DOCUMENT RESUME

ED 096 348 TM 003 955

AUTHOR	Bryant, Edward C.; And Others
TITLE	Associations Between Educational Outcomes and Background Variables: A Review of Selected Literature.
INSTITUTION	Westat Research, Inc., Rockville, Md.
SPONS AGENCY	Education Commission of the States, Denver, Colo. National Assessment of Educational Progress.
PUB DATE	74
NOTE	216p.; For a related document, see TM 003 956
AVAILABLE FROM	Westat, Inc., 11600 Nebel Street, Rockville, Maryland 20852 ($4.00)
EDRS PRICE	MF-$0.75 HC-$10.20 PLUS POSTAGE
DESCRIPTORS	*Academic Achievement; Aspiration; *Background; *Data Analysis; Education; Educational Assessment; *Environmental Influences; Expectation; Income; *Literature Reviews; Motivation; Occupations; Performance Factors; Relationship; School Environment; Self Concept; Socioeconomic Status; Student Attitudes; Student Characteristics
IDENTIFIERS	*Educational Outcomes; National Assessment of Educational Progress

ABSTRACT

The literature is reviewed that addresses the association between background factors and educational outcomes. After completing the review, Westat, Inc. issued this report which synthesizes the literature into correlates of academic performance and outcomes other than academic achievement. The former considers such background variables as socioeconomic status, personal characteristics of students, school variables, attitudes, motivations, self-perception, aspirations, intentions, and expectations. The latter discusses the relation of background factors to education, occupation, and income; considers socioeconomic status, ability, and attendance at college; looks at job experience as an outcome of education; and studies attitudes and motivations. Summaries of both sections appear, and a section detailing the report's conclusions and an epilogue describing the methodology used in relating background variables to outcomes are included. (SE)

ED 096348

ASSOCIATIONS BETWEEN EDUCATIONAL OUTCOMES and BACKGROUND VARIABLES:

A Review of Selected Literature

by

Edward C. Bryant
Ezra Glaser
Morris H. Hansen
Arthur Kirsch

of

Westat, Inc.

Under contract to

Education Commission of the States
300 Lincoln Tower
Denver, Colorado

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM 003 955

A Monograph of the National Assessment of Educational Progress

This publication was prepared and produced pursuant to agreements with the National Center for Educational Statistics of the U.S. Office of Education with additional funds from the Carnegie Corporation of New York and the Ford Foundation's Fund for the Advancement of Education. The statements and views expressed herein do not necessarily reflect the position and policy of the U.S. Office of Education or other grantors but are solely the responsibility of the National Assessment of Educational Progress, a project of the Education Commission of the States.

National Assessment of Educational Progress, 1974

Single copy $4

Westat, Inc.
11600 Nebel Street
Rockville, MD 20852

National Assessment of
Educational Progress
Suite 700, 1860 Lincoln
Denver, CO 80203

TABLE OF CONTENTS

FOREWORD. v

PREFACE ix

Chapter Page

1 INTRODUCTION. 1

1.1 Purpose of the Study 1
1.2 Some Definitional Matters. . . . 3
1.3 A Note on Measurement of Association. 6
1.4 An Application of an Association Model to NAEP Exercises. 14

2 CORRELATES OF ACADEMIC PERFORMANCE. . 21

2.1 Overview 21
2.2 Measures of Socioeconomic Status (SES). 22
2.3 Personal Characteristics of Students 57
2.4 School Variables 75
2.5 Other Background Factors, Including Attitudes, Motivations, Self-Perception, Aspirations, Intentions, and Expectations . . 115
2.6 Summary. 131

3 OUTCOMES OTHER THAN ACADEMIC ACHIEVEMENT 135

3.1 Overview and Summary 135
3.2 Relation of background factors to education, occupation, and income 139
3.3 SES, Ability, and Attendance at College 145
3.4 Job Experience as an Outcome of Education. 147
3.5 Attitudes and Motivations. . . . 163
3.6 Summary of Findings for Outcomes Other Than Achievement 177

TABLE OF CONTENTS (Continued)

Chapter Page

4 CONCLUSIONS 177

4.1 Measures of SES. 177
4.2 Sex. 180
4.3 Race 180
4.4 Ability. 181
4.5 School Factors 181
4.6 Motivation, Attitude, Expectation and Other Measures. 181

5 EPILOGUE. 183

5.1 Introduction 183
5.2 Methodology for Estimating Association. 183
5.3 Sources of Errors. 187
5.4 Limitations in the Interpretation of Relationships from Survey Data 188
5.5 Unit of Observation and Analysis 190
5.6 Large-Study Bias 190
5.7 Background Adjustment of Longitudinal Data 191
5.8 Some Additional Comments and Recommendations Concerning Choices of Background Variables for NAEP 194
5.9 Some Additional Suggestions for Research 197

CITED REFERENCES. 201

FOREWORD

The purpose of the National Assessment of Educational Progress is to chart changes in American education by periodically measuring the knowledge, skills and attitudes of 9-year-olds, 13-year-olds, and 17-year-olds and young adults (26-35). In particular, National Assessment estimates the proportion of individuals who can respond acceptably to certain questions or tasks (called exercises) in Science,, Mathematics, Reading, Literature, Writing, Citizenship, Social Studies, Art, Music and Career and Occupational Development.

Traditionally, several variables -- more accurately, "bundles" of variables that are proxies for each other -- have been associated with differences in educational achievement in this country. National Assessment analyses have employed several of these traditional variables in order to group respondents at each age level into subpopulations. These variables include region of the country, sex, color, level of parents' education, and size and type of community. National Assessment reports present the percentages of people in each subpopulation performing acceptably on an exercise and they compare these group percentages to the percent of all persons in the age population performing acceptably. The difference between these two figures is called the group effect.

For some time, National Assessment has been concerned that there may be other reporting variables that are better than these -- that is, there may be more sensitive variables with fewer proxies. Some may be more easily and reliably measured; others may be more "durable" over time; some may be more relevant to certain subject areas or ages; and others of greater use for interpretation and policy formation in the educational community.

Another National Assessment concern is how to better understand and interpret group effects. One can perform certain statistical data adjustments to gain a better understanding of the facts obtained from our surveys. We have already employed one adjustment called "balancing." The fact that a group effect reflects Northeast or Southeast regional performances does not mean that these performances occur solely because the respondents live in the Northeast or Southeast. Since a larger fraction of respondents in large cities live in the Northeast than in the Southeast, and since a larger fraction of respondents in rural areas lives in the Southeast than in the Northeast, regional differences may be masquerading as size and type of community effects. Balancing is intended to remove this masquerading and avoid double counting when making comparisons.

Any adjustment procedure, however, is heavily dependent on the nature and number of variables included in the adjustment. The fact that proxy bundles for variables may change over time and that the representation of one group in another is also changing makes adjustment of measures of change in performance even more complex.

These and similar concerns led to the decision that National Assessment identify further variables useful in gaining a better understanding of differences in American education. Subsequently, a contract was let to Westat, Inc. to survey relevant papers and sources purporting to show relationships between background variables and educational outcomes.

The specific objective given to Westat in undertaking the literature search was to catalog past and current research into important background factors. The search was to include studies undertaken by scholars in disciplines other than education and in a variety of contexts other than the school. Further, the Westat report was to describe the methods used by various researchers to identify, delineate and measure background variables.

In addition to this information about important studies, National Assessment asked for a bibliography of studies identified by Westat's literature search but not critiqued and detailed. Such a bibliography would be useful not only to National Assessment but to other users of this background factors study who may wish to initiate further research.

With this background factors study, Westat has concluded the first of many important steps toward the possible adoption of new and more useful variables for National Assessment analyses and reports.

Roger Talle
Robert Larson

PREFACE

This report was prepared under contract to the Education Commission of the States and was supervised by Dr. Roger Talle, Director of Operations for the National Assessment of Educational Progress (NAEP). The purposes of the contract are identified in an "Introduction to Background Variables Report" which follows this Preface.

This report has a strong statistical flavor. As a consequence, there is a minimum of interpretation of the results in terms of educational policy, although the relevance to educational policy of the investigated relationships is obvious. We have purposely chosen to let others draw such policy relevance from the conclusions reported.

It would not be feasible for us to acknowledge here all of the persons who contributed to the completion of this report. A name that must be mentioned, however, is that of Dr. Torsten Husén, Chairman of the International Project for the Evaluation of Educational Achievement, who provided invaluable advice in the planning stages of the project. Another major contribution was that of James Gold and Gordon Hanson, compilers of Correlates of Achievement, a helpful draft bibliography. Dr. Carl Feigenbaum of the Westat staff contributed to the project in a major way by his painstaking review and screening of many of the items included in our bibliography (separately bound).

As a part of the project, we invited the participation of a number of educators, representatives of state education departments, educational researchers, and government officials. The preliminary draft was submitted to them and a two-day critique session was held at the Rockville offices of Westat on December 5-6, 1973, with approximately half of the participants at each day's session. The participants and their affiliations are listed on the following pages.

John W. Adams
Director, State Educational
Assessment
Minnesota Department of
Education
731 Capitol Square Building
St. Paul, Minn. 55101

Jay Davis
Director, Center for Educational
Research and Evaluation
Research Triangle Institute
Research Triangle Park, NC 27709

Dorothy Gilford
Assistant Commissioner for Educational Statistics
U.S. Office of Education
400 Maryland Avenue, S.W.,
Room 3073
Washington, D.C. 20202

Thomas E. Kendig
Chief, Division of Educational
Quality Assessment
Pennsylvania Department of
Education
P.O. Box 911
Harrisburg, Pa. 17126

Garry L. McDaniels
Chairman, Educational Personnel
Task Force
National Institute of Education
Room 815, Brown Building
1200 19th Street, N.W.
Washington, D.C. 20208

George W. Mayeske
Research Psychologist
Office of Planning, Budget and
Evaluation
U.S. Office of Education
400 Maryland Avenue, S.W.
Washington, D.C. 20202

Selma J. Mushkin
Director, Public Services
Laboratory
Georgetown University
Washington, D.C. 20007

Cyril B. Busbee
State Superintendent of Education,
South Carolina
State Department of Education of
South Carolina
Education Commission of the
States, Steering Committee
1429 Senate Street
Columbia, S.C. 29201

Iris Garfield
Project Office, National Assessment Educational Progress
National Center for Educational
Statistics
U.S. Office of Education
400 Maryland Avenue, S.W.
Washington, D.C. 20202

Thomas L. Hilton
Senior Research Psychologist
Educational Testing Service
Rosedale Road
Princeton, N.J. 08540

Frank W. Kovacs
Director of Program Evaluation
National Education Association
1201 16th Street, N.W.
Washington, D.C. 20036

James McPartland
Assistant Director
Center for Social Organization
of Schools
Johns Hopkins University
Baltimore, Md. 21218

Mary Milne
Policy Analyst
U.S. Office of Education
400 Maryland Avenue, S.W.
Washington, D.C. 20202

Andrew C. Porter
Visiting Scholar
National Institute of Education
1200 19th Street, N.W.
Washington, D.C. 20008

Marshall S. Smith
Visiting Scholar
National Institute of Education
1200 19th Street, N.W.
Washington, D.C. 20202

Tongsoo Song
Education Program Specialist
National Center for Educational Statistics
U.S. Office of Education
400 Maryland Avenue, S.W.
Washington, D.C. 20202

Judith V. Torney
Associate Professor of Psychology and Education
University of Illinois at Chicago Circle
Chicago, Ill. 60680

Nicholas Zill
Staff Associate
Social Science Research Council
Center for Social Indicators
1785 Massachusetts Avenue, N.W.
Washington, D.C. 20036

The participation of this distinguished group contributed in a major way to the report. While we have attempted to reflect the comments of the critique panel in our final report, it must not be inferred that the report has been endorsed by the panel. The authors alone take the responsibility for the selection of literature to be reviewed and the interpretation of results.

Edward C. Bryant
Project Director

CHAPTER 1

INTRODUCTION

1.1 Purpose of the Study

Within the past ten years substantial interest has developed in the measurement of educational outcomes and their relation to various inputs. Some of the interest has stemmed from the desire to evaluate the effectiveness of educational processes, both with regard to particular strategies and also with regard to the total milieu within which education takes place. The National Assessment of Educational Progress (NAEP) is a long-range project of this kind. A special interest in educational outcomes has been generated by concern about equality of educational opportunity for specific minority groups. At one extreme of this concern is the philosophy that equality of educational opportunity is evidenced by equality of educational outcomes. A less extreme view is concerned with improving educational strategies and recognizes the importance of the measurement of outcomes in the evaluation of alternatives.

It also has been observed that measures of educational outcomes vary, not only by the educational process applied to the pupils, but also by the socioeconomic backgrounds and other characteristics of the pupils -- their family composition, the income and educational attainment of their parents, and so on -- as well as by the ability of the pupils. This recognition has led to the use of background measures to adjust outcome measures (by statistical procedures) in order to permit more sensitive comparison of outcomes where such outcomes are intended to reflect effects of educational processes or of efforts to equalize educational opportunity.

NAEP is doubly interested in adjustment for background factors. First, it is interested in comparing the outcomes of subgroups of the population. The absolute differences of those subgroups are of interest to educators, sociologists, and government planners because they represent the gross sum of all influences on the population subgroups, and gross differences may provide sufficient evidence of the need for programmatic action. But educators and others are also interested in subgroup differences, net of differences in background, as a way to measure the effectiveness of the educational process. An example of this use of data adjustment is provided by NAEP's Report 7 for the 1969-70 assessment which provides "balanced" results for science.

Second, by its name, NAEP is interested in measuring the *progress* of education, implying comparisons over time. Thus, it is important to adjust apparent differences over time for differences in backgrounds that are known to have an impact on educational outcomes. Otherwise, apparent changes over time could be accounted for by a changing composition of the sample, which may reflect a change in the composition of the population as well.

In this report we have examined some of the principal literature which shows association between background factors and measures of educational outcome. We have not attempted to examine *all* such literature. An extensive bibliography containing several hundred items is included as an appendix to this report and is bound separately. Out of those items we have selected some which appear to contain the principal conclusions that are portrayed by the larger set. Other researchers might have chosen a different set and, indeed, the bibliography might easily have been doubled or tripled in size by devoting more time to the search. Some of the principal papers cited in this report have been abstracted and the abstracts have been included in the bibliography.

One of the findings of nearly all researchers in the field is that background factors are highly intercorrelated. Education of parents is correlated with occupations of parents and both are correlated with family income. If one adjusts outcome measures for differences in educational attainment of parents he will also adjust for much of the difference in occupation (or income). Thus, there is a great potential for substitution among the background variables measured. We have only incidentally noted presumed difficulty or ease of measurement of background variables, although this characteristic would certainly be considered in choice of variables in any revision of the NAEP data-collection plan. Intercorrelations among variables are important, of course, to the feasibility of selecting variables at least partly on the basis of ease in collection.

Outcome variables are also correlated, and the literature is filled with research on testing that shows relationships among measures of academic performance, attitudes, aspirations, job performance, and so on. We have not given much attention to such correlations in this report. Instead, we have organized the report by principal classes of outcome measures. Chapter 2 discusses academic achievement scores and related kinds of outcome measures and the relationship of these to various types of background variables. Socioeconomic background variables are discussed first, followed by personal characteristics and family relationships, then by school variables and finally by attitudes, motivations, self-perception, self-esteem, external control, aspirations, intentions and expectations, where this last group of variables is considered as background for achievement rather than outcome of the educational process. Chapter 3 discusses other outcomes briefly, such as attitudes, social participation, college attendance, and so on. Chapter 4 summarizes the principal findings and Chapter 5 presents some of the staff's views on related topics.

In the remainder of Chapter 1 we briefly discuss some definitions and methodology that affect the interpretation of the findings presented in Chapters 2 and 3.

1.2 Some Definitional Matters

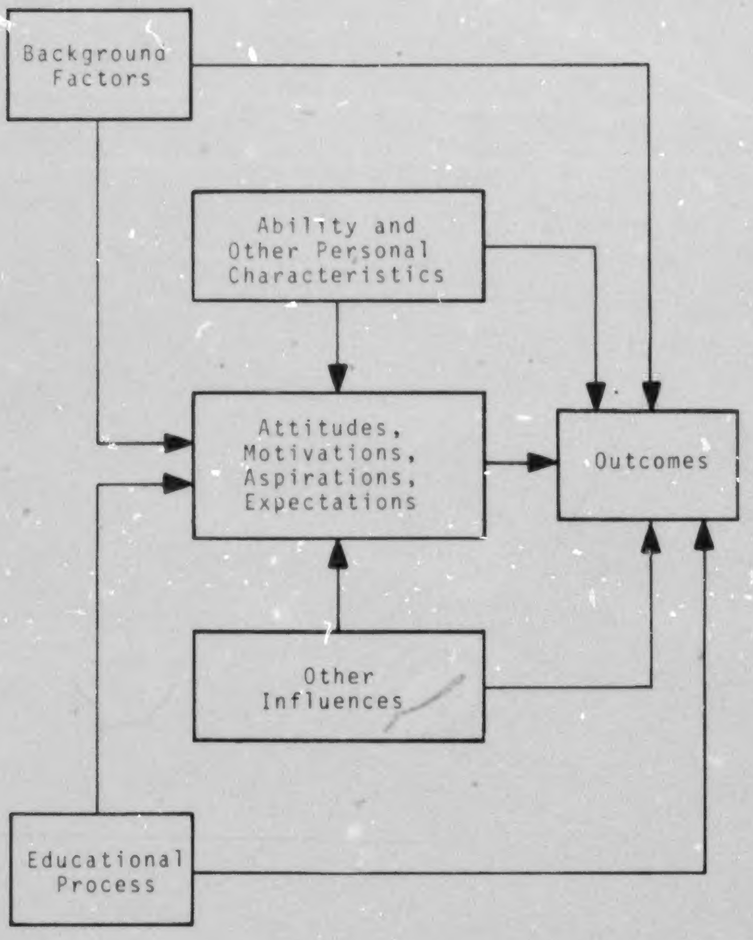
The purpose of this study is to review the principal literature that addresses the association between background factors and educational outcomes. The intended use of the study is to provide the basis for decisions concerning the kinds of background data to be obtained from pupils, teachers, school administrators, and other sources relating to assessment and analysis of educational progress. The technique used is to identify the fraction of the variability of a measure of outcome that can be associated with variability in the background variables, jointly or separately, and thereby to help explain the variability in the outcome measures. The principal outcomes of interest are the outcomes measured by the National Assessment. These consist primarily of answers to exercises, the percent of answers of a given kind by a given population subgroup being a measure of accumulated knowledge, skill, or attitude of the subgroup. It can be anticipated that performance on such a set of exercises will be correlated with performance on more traditional kinds of educational tests.

The term "educational outcomes" is not intended to imply that the outcome is the result of only the formal educational process. Knowledge, skills, and attitudes are influenced by numerous other factors, some of which are measurable or correlated with things that are measurable but many of which cannot be calibrated in any reasonably adequate manner.

The measurement of outcomes and their comparison over time may permit one to "assess" educational progress (where "education" is broadly construed to include informal processes), but without additional information they are of limited use for analytical purposes. For interpretation one also wants to know how effective the total educational process has been, both nationally and for specified subgroups, after allowance for differences that tend to handicap or give advantage to specified subgroups of interest. This analytical requirement provides an important motivation for collecting background variables.

Some background factors are relatively stable, such as the occupation and education of parents, and are not influenced by the educational process. It is helpful in evaluating the educational process to distinguish these from others that may be controlled or changed. The key assumption is that, for the short term, some background variables may be considered fixed, and these, especially, may be regarded as not subject to the control of the educational process. A simplified model has been displayed in Figure 1. The arrows may be interpreted to symbolize the words "has an influence on." The model follows in a general way the "path coefficient" models of Duncan, et al. (11)

Figure 1. Educational Associations Model


Background Factors
Ability and Other Personal Characteristics
Attitudes, Motivations, Aspirations, Expectations
Outcomes
Other Influences
Educational Process

For these purposes we have assumed that a measure of ability is a relatively stable background factor. For analytical purposes it may be desirable to isolate it from other background variables. It is assumed that ability is not influenced by the educational process (although there is ample evidence that the measurement of ability may be).

Note particularly, however, that motivations, expectations, aspirations, and other attitudes may be influenced both by the educational process and by more stable background factors, and in turn, exert an influence on outcomes. Thus, they may be considered intermediate variables. For some purposes it is desirable to consider them as outputs and for others to consider them as inputs. In order to keep this study within manageable size we have deemphasized intermediate variables and have generally studied reported associations between more stable background factors and performance.

1.3 A Note on Measurement of Association

Suppose we consider a simple (but hypothetical) example in which the outcome is test score on a twelfth-grade science test and the background variables are race and sex. We assume that such a large sample of twelfth-grade students has been tested that we need not be concerned about sampling error. Suppose the averages are as follows (the data are hypothetical):

	Race		
Sex	White	All other	All Races
Male	64	56	63
Female	60	52	59
Both sexes	62	54	61

There are obvious differences in the average performance of the sexes and the two categories of race, because they vary from the over-all average score of 61.

If a male white student achieves a score of 70, one can identify 70-61, or 9 score points, as the amount of his total variation from the mean of 61. Of that total variation of 9 points, three points (64-61) are "explained" by the fact that he is a white male, and the remainder (6 points) is unexplained by the sex-race classification.

The above example is an illustration of a means of explaining some of the sources of variation in an individual student's score. A way is needed to summarize such variation over all students. Because of properties that need not be discussed here, sums of squares of difference are used. It can be shown that:

A. The sum of the squared differences between individual-student scores and the overall mean is equal to:

B. The sum of the squared differences between the individual scores and the individual cell (sex by race) means, plus

C. The sum of the squared differences between the cell means and the over-all mean (summed over all students).

The ratio (C/A) is the proportion of variance explained by the background variables -- in the above illustration, by sex and race.

In the real world one does not have test scores for all students in the universe, so there is sampling error in the estimation of the means. Also, even if he had scores for everyone, there would be measurement error that would tend to distort the means. Such sampling and measurement errors tend to make estimation of the proportion of explained variance less precise, but the additive relationship given above still holds.

Note that "explained" does not connote cause and effect, only association. Also, it may be that race (for example) is strongly associated with income and other factors, and these factors might have explained more variance than race if they had been used.

The introduction of many background variables together with relatively small samples makes the simple analysis illustrated above infeasible -- there are too many groups (cells in the table), and some of them may have no cases in them. Two procedures are commonly used to approximate the above type of analysis. First, multiple linear regression methods are employed to estimate the proportion accounted for by the background factors. That is, in the above illustration, one could set up the following linear model:

$$y_i = Sx_{1i} + Rx_{2i} + Ix_{1i}x_{2i} + e_i \quad (i = 1, 2 \ldots n)$$

where

n = the total number of students,
x_1 = 1 if male, 0 if female
x_2 = 1 if white, 0 if nonwhite
y = score in science
S = effect of sex
R = effect of race
I = interaction effect of race and sex
e = unexplained variation

In the above example S and R are *marginal effects* and I is an interaction. In this simple case, solving for S, R and I by least squares methods and using those results to estimate the proportion of variance explained by sex and race would yield the identical results obtained by examining the cell means, above.[1] The equivalence will not be shown here. An important feature of linear regression is that it is extendable to many variables. However, if there were three marginal effects one could have three simple interactions (among two factors) and one three-factor interaction. The number of interactions expands rapidly with increasing numbers of factors and an almost universal practice is to assume higher-ordered interactions (among three or more factors) are zero and hence are not to be computed. Also variables that are found not to have an important effect, or that essentially duplicate the effects of other variables, are eliminated. The result is a simplification that permits one to estimate the proportion of variation explained by a large number of variables.

---

[1] This is not true, in general, for tables with more than four cells.

The second procedure commonly used to reduce the number of variables (or the number of groups for which effects are to be computed) is to construct composites of variables. For example, Thorndike (48, pp. 72ff), in his analysis of reading-comprehension scores in the U.S. portion of the international assessment, used ten categories for father's occupation. These ten classes were used as variables in a regression analysis, each receiving a value of one for the category in which father's occupation fell and zero otherwise. The criterion variable (dependent variable) was reading-comprehension score for the 14-year-old population. The regression analysis automatically assigned weights to each occupational category and thus collapsed the ten categories into a single score for father's occupation. Scales were also developed for father's educational attainment and mother's educational attainment. The three scales were put into a single composite measure of SES, using similar procedures. The method is quite general and has been used extensively by analysts of educational test scores. A procedure such as that just described would lead to a different weighted average, or composite, for each dependent variable -- e.g., one composite measure for reading and another for mathematics scores. A common procedure, however, is to define a single composite to be used with a number of dependent variables being analyzed.

Another method of forming composite scales is to use factor-analytic methods and to identify the principal factors by the nature of the individual scores that comprise them. The method will not be discussed here.

There is still another kind of compositing or averaging that has a major effect on the amount of explained variance. Sometimes the student is the unit of analysis, and sometimes the school is the unit of analysis. In the former case one is interested in determining the percent of variation in student outcome measures explained by background variables, and in the latter case one is interested in determining the percent of variation in *average school scores* accounted for by background variables *averaged over the school*. The percent explained when the school is the unit of analysis is generally higher than when the student is the unit of analysis. An example of the importance of the unit of analysis in estimating percent of variance explained is provided by Comber, et al. (7) with respect to the U.S. scores on the International Studies in Science Education. Four composites of background factors were constructed by methods similar to those described above. The following data represent percentages of variance explained for 14-year-old students.

	Additional percent explained	
Composite variable	School as unit	Student as unit
SES, sex	67	22
Type of school	0	2
Learning conditions	11	7
Kindred variables	1	6
Total	79	36

The total percentage explained is approximately half as great for students as for schools, and SES serves as a much more effective explanation of differences in average school scores than of differences in student scores. These results are not atypical. Note that variance of student scores is across all schools and not variance within schools.

Display of the components of total sum of squares (of individual student scores from their overall mean) helps one to understand the difference between variance explained by regression on school means and variance explained by regression on individual student scores. Consider the following sources of variation:

Source of variation	Sum of Square	
Due to differences in school means		A
Due to regression on school means	A_1	

Source of variation	Sum of Square	
Residual from regression on school means	A_2	
Due to differences among students within schools		B
Due to common regression on student scores within schools	B_1	
Residual from common regression within schools	B_2	
Total sum of squares		C

The above display assumes that there are two regressions: one on school means and the other on student scores within schools. Although this kind of analysis would be helpful in determining the impact of certain background variables, it has not been widely used. A commonly used analysis is the following:

Source of variation	Sum of Square
Due to regression on student scores, ignoring school differences	C_1
Residual from student regression	C_2
Total sum of squares	C

The percent of variation explained by regression on *individual student scores*, ignoring differences among schools, is $100C_1/C$. The percent of variation explained by regression on *school means* (from the previous model) is $100A_1/A$.

The differences in magnitude of the percents of explained variance can be analyzed in the following way.

Let R_b^2 = the proportion of variance among school means of achievement scores accounted for by regression on the school means of the regressor variables = A_1/A.

R_w^2 = the proportion of variance within schools accounted for by regression on the individual student regressor variables after subtraction of the school means = B_1/B.

p = proportion of total student variation in achievement scores that lies between school means = A/C.

R_{b+w}^2 = the proportion of total student achievement variance accounted for by the two regressions (one on school means and one on student characteristics after subtraction of school means) = $(A_1 + B_1)/C$.

R^2 = the proportion of total student achievement variance accounted for by a single regression (ignoring differences among school means) = C_1/C.

Then,

$$R_{b+w}^2 = p\, R_b^2 + (1 - p)\, R_w^2$$

We have seen that R_b^2 is typically an overestimate of R^2. Ordinarily, $R_b^2 > R_{b+w}^2 > R^2$. In the NAEP applications one would like to know R^2, which is derived from a single regression equation. It tends to be only slightly smaller than R_{b+w}^2, but R_b^2, the variance explained by regression on school means, tends to be substantially higher than R^2.

An important but often ignored characteristic of regression and correlation studies is that the independent variables chosen for the regression are those that show high correlation in the particular sample observed. Some of this apparent high correlation is the consequence of that particular sample -- some variables would have higher and some lower correlations with the criterion in another sample of similar size and design. Unless the samples are very large, this choice of variables to include in the regression equation will tend to overstate the "true" correlations and the "true" explained variance. A computation of the same correlation (based on the same independent variables) in an independently selected sample typically would yield a lower and more valid estimate of the correlation. Thus, regression and correlation computations tend to result in overstatements of the amount of explained variance. An illustration of this is given in a California State Department of Education study (4). After a set of independent variables was selected through step-wise regression, correlations were estimated from each of a set of dependent variables. The median squared multiple correlation coefficient was approximately 0.50. However, a separate sample had been withheld from the analysis employed in choosing the independent variables. When the correlation was estimated from this independent sample, the median squared multiple correlation coefficient was reduced to approximately 0.40. Ordinarily, no such independent sample is withheld for evaluation and estimation. Indeed, for this reason, most of the analyses included in this report can be regarded as overestimates of correlations or of explained variance. The magnitude of the problem is a function of the size of the sample -- with quite large samples there is less danger of substantial overestimation. In the California study, school districts were the unit of analysis, and the regression estimates involved the selection of 11 independent variables for use in a regression equation from a set of 22 that were examined.

Another important characteristic of regression and correlation studies is that, as variables are added to a regression equation, it is common to show the increase in the multiple correlation as a result of the added variable, and the increase in the amount of variance accounted for. Thus, the added amount of variance accounted for by a variable is a function of the order in which it is introduced and of the other variables that have been introduced. A different ordering of the variables may assign them quite different proportions of variance accounted for. The effects of this are observed in many of the studies we have examined. It is common, for example, to introduce a measure of SES early, and consequently this will account for a higher proportion of the variance than if school variables were introduced and the socioeconomic-status variable later. Unless proper account is taken of this some of the results cited will appear contradictory when they may not be.

One more general comment is appropriate. The efforts to adjust for various measures can be effective only to the extent that the measures for which adjustment is being made are themselves reasonably accurate measures. The effectiveness of adjustments is reduced when independent variables which are themselves subject to substantial errors of measurement are used for adjustment. The practical consequence is that adjustment for variables like sex will not be significantly impaired by errors of measurement, but adjustments for socioeconomic status, or measures of ability, or other variables that are subject to substantial errors of measurement may be impaired.

1.4 An Application of an Association Model to NAEP Exercises

NAEP administers exercises to 9, 13, and 17-year olds and to young adults. While the exercises are grouped according to subject matter into packages for administration, aggregate test scores are not computed for individuals. The exercise itself is the focus for analytical attention and aggregation is accomplished for a single exercise across all students by computing "P values," i.e., the proportion of all answers falling into a given category. Usually it is possible to identify one of the response categories as correct. In such cases the P value for that response category is the percent of students responding correctly. In some cases, such as in attitude exercises, there is not a correct category. In those cases the P values for all of the response categories must be examined together.

In order to focus attention more sharply on what is meant by "association" and "data adjustment" we consider the models given below. For convenience, we consider the case in which there is an identified correct response to the exercise. Then, for the e^{th} exercise, the j^{th} member of the i^{th} population subgroup will achieve a score of 1 if he answers correctly or 0 if incorrectly, which we will designate by the variable y_{eij}. This variable is the outcome measure for the j^{th} member of the i^{th} subgroup attempting the e^{th} exercise. We presume that this outcome is a function of background effects (which may be different for each exercise and for each subgroup) and unexplained factors, so that we may specify:

$$y_{eij} = f_e \, (F_{eij}, I_{eij}, S_{eij}) + E_{eij}$$

where F_{eij} is a set of family background characteristics associated with the correctness of response to the e^{th} exercise by the j^{th} member of the i^{th} subgroup, I_{eij} is a set of individual student characteristics such as sex, ability and motivation, S_{eij} is a set of school characteristics (including characteristics of the community in which the school is located). The component E_{eij} is an unexplained component of the response which includes response error, failure to include explanatory variables, and random error. The functional form of f_e is unspecified, but in actual practice a linear model has usually been presumed. Also, in practice, the motivational and attitudinal characteristics are frequently included in the term E_{eij}, rather than in I_{eij}. If one averages actual responses over the i^{th} subgroup for the exercise, he obtains P values.

Computationally, f_e is often considered to be a linear function of family, individual and school characteristics. In a simple case, one could construct three indices of background factors: one for family characteristics, one for individual characteristics, and one for school characteristics. If only one index is used for an F-type characteristic, one for an I-type characteristic, and one for an S-type, a simplified model could be expressed as:

$$p_{eij} = b_{0e} + b_{1e} F_{ij} + b_{2e} I_{ij} + b_{3e} S_{ij}$$

where p_{eij} can be regarded as a regression estimate of the probability that individual j in the i^{th} group will answer exercise e correctly. It will be a close estimate if the multiple correlation coefficient is near unity. Obviously, the model can be expanded to include cross-products of the indices or to include individual variables for each of the three types and transformations of them.

Using the simplified model, above, the total sum of squares over all subgroups for exercise e can be expressed as:

$$\text{Total SS} = \sum_{i,j} (y_{eij} - P_e)^2$$

where y_{eij} is the actual score (zero or one) and P_e is the average over all i and j. The sum of squares accounted for by the indices F, I and S can be expressed as:

$$\underset{\text{(by F, I and S)}}{\text{SS Accounted for}} = \sum_{i,j} (p_{eij} - P_e)^2$$

Then, the proportion of the total variation in individual exercise responses accounted for by the regression model is

$$R^2 = \frac{\text{Accounted for SS}}{\text{Total SS}}$$

and this is the square of the multiple correlation coefficient.

To illustrate the adjustment of data with the above procedure for a particular exercise, e, let i = 1, 2 designate two subgroups, black and other, for which comparisons are to be made, and j the individual student taking the exercise within the i^{th} group. Suppose that F_{ij} is a composite measure that represents the socioeconomic status based on a composite measure of social status of the family of student ij. For illustration, F_{ij} might take on values as follows:

1 = high measure of socioeconomic status
2 = middle measure of socioeconomic status
3 = low measure of socioeconomic status

Let I_{ij} be assigned the values 0 or 1 depending on whether the student designated by ij is male or female, respectively, and let S_{ij} represent the average expenditures per pupil in the school attended by student ij. The observed score on a particular exercise for group i is:

$$p_{ei} = \sum_{j}^{n_i} y_{eij}/n_i$$

and the score that would be produced from the model if one assumes a "common regression" across all subgroups is:

$$p'_{ei} = b_{0e} + b_{1e}\,\bar{F}_i + b_{2e}\,\bar{I}_i + b_{3e}\,\bar{S}_i$$

where $\bar{F}_i$, $\bar{I}_i$ and $\bar{S}_i$ are averages of the background scores for the i^{th} group.

Let $\bar{F}$, $\bar{I}$ and $\bar{S}$ represent average background scores for all students taking exercise e. Then, an adjusted P value for group i is:

$$\hat{p}_{ei} = p_{ei} + b_{1e}\,(\bar{F} - \bar{F}_i) + b_{2e}\,(\bar{I} - \bar{I}_i)$$

$$+\, b_{3e}\,(\bar{S} - \bar{S}_i)$$

This adjusted score will, *to the extent that the model is effective in doing so*, remove the effects of differences among the subgroups for the variable used in the model; i.e., it will remove the effects of differences in proportions in different socioeconomic classes, the differences in the proportions of each sex, and the differences in the levels of expenditures per student.

There is an assumption in the model that the regression on the background characteristics is the same for all subgroups of interest. This assumption frequently is not true. For example the regression of performance on SFS may be different for black children than for white children. In such cases, if one wishes to compare black children with white children the following adjustment procedures are sometimes employed:

1. Compute the regression of white outcome on white SFS, and arrive at SFS-adjusted P values.

2. Apply those regression coefficients to the black background measures and arrive at SFS-adjusted P values for black children.

3. Compare the adjusted P values.

This procedure essentially matches that employed by Duncan, et al.(11). The risks in such adjustments are evident. The regressions are likely to be different in the two groups so that, if the whites were adjusted by the black regression, a different comparison would be obtained.

Another approach is to adjust each group by its within-group regression, as follows:

$$p_{ei} = p_{ei} + b_{1ei}\ (\bar{F} - \bar{F}_i) + b_{2ei}\ (\bar{I} - \bar{I}_i) + b_{3ei}\ (\bar{S} - \bar{S}_i)$$

That is, a separate regression adjustment might be made for each group of interest, and the adjusted scores can then be compared.

Another adjustment follows the procedure, used in demographic computations, of computing adjusted rates. Mushkin (38) suggests three methods of adjustment, based upon weighting to population values, that are essentially of this form. There are, of course, other adjustment procedures. The ones we have described are common ones. The one used by NAEP in Report 7, which balances science-exercise P values, is less common.

CHAPTER 2

CORRELATES OF ACADEMIC PERFORMANCE

2.1 Overview

This chapter discusses the relationships that have been examined in the literature between background factors and measures of academic performance. NAEP's measure of outcome is a response to an exercise which is designed to measure knowledge, skill or attitude with respect to an educational objective. The outcomes in this chapter can most easily be identified with "knowledge."

Most of the literature we have examined considers outcomes in the form of test scores rather than responses to individual exercises (as in NAEP), but the assumption that the association with background factors would be similar seems reasonable. Also, some of the most complete analyses of associations with background variables have used measures of ability, rather than performance, as the criterion. We have used some of these studies to show associations under the assumption that many such ability measures are highly correlated with performance and many, in fact, have strong components of accumulated knowledge in them.

In Section 2.2 of this chapter we discuss the association between socioeconomic status (SES) and outcomes. SES is reasonably independent in the short run of school influences.

In Section 2.3 we discuss the association between outcomes and personal characteristics, family structure, student-parent relationships and measures of ability. In particular, we have looked for research papers that have distinguished the added amount of association due to ability over the association due to home background and other personal characteristics. Note that in Section 2.2 we sometimes use ability as a proxy for academic performance, but in Section 2.3 we consider individual ability as a background variable.

Section 2.4 considers school variables as background in the explanation of academic performance. We are primarily interested in determining how much impact school characteristics have on individual performance. Unfortunately, much of the literature deals with the predictability of *average school outcomes* from school variables, including *average student-body characteristics*. It has not been possible to unravel most of the published figures in a way that would be optimally useful to NAEP.

Section 2.5 considers the impact of other background factors on educational performance. Variables considered are attitudes, motivations, self-perception, aspirations, intentions, expectations, and other factors. They frequently have been called "intermediate" variables since for some purposes they are an output of the educational system and for other purposes they are an input. We consider them as inputs in Chapter 2 and as outputs in Chapter 3.

2.2 Measures of Socioeconomic Status (SES)

One set of background factors that is shown by nearly every study of background to have relatively high association with measures of school performance is socioeconomic status. This generalization holds over a wide range of outcome measures and over a variety of ways in which SES is measured. It also holds for both the performance of schools and for the performance of *individuals within schools*. Therefore, we have chosen to begin our discussion of background with variables that measure SES and to presume for subsequent analysis that some measure of SES will be included in any measure of background.

One is immediately faced with the reality that the various measures of SES, as well as various outcome measures, are nontrivially correlated. That is, if measures of SES are associated with scores in arithmetic they are likely also to be correlated with scores in reading, in languages, in measures of ability and a wide assortment of affective measures. Also, various measures of SES, such as family income, parental education, occupation of household head, rental value, reading materials in the home and appliances in the home are highly intercorrelated.

Consequently, using a single carefully chosen measure of SES, such as family income or occupation of head of household, may "explain" a large proportion of the variance associated with measures of SES, and adding additional measures will produce smaller and smaller increments of explained variance. Furthermore, this generalization is reasonably invariant for most academic outcomes. There is less stability over the various affective measures, but even there a substantial amount of invariance exists.

Since there is extensive evidence of the importance of SES to educational outcomes we will accept the need for its measurement. The thrust of this part of the present study will be to examine association between composite measures of SFS and educational performance and then to investigate the components of various SES measures.

Although some studies have included race or national origin in SES, we assume, because there has been so much interest in recent years in equalizing educational opportunity, that race or ethnic-group membership will be recorded as a separate item and not be included in a composite measure of SES. In this study race or ethnicity are separated from measures of SES.

It seems clear that a measure of family income or a proxy for it should be included in a socioeconomic measure. Commonly used proxies are occupation of father (and sometimes the mother) and items in the home. Measures of social status frequently have been defined in terms of personal and family characteristics which in some way are presumed to have an association (not necessarily causative) with educational outcomes. One such measure is education of parents, although the correlations of this item with occupation of head of household (and with income) are quite high. A number of studies have tended to build social status into a classification of occupations (e.g., 27, pp. 19ff) rather than simply economic level. The NORC study (Reiss (47)) made a pioneering effort to associate "general standing" with a wide range of occupations.

It is common for socioeconomic scales to include measures that are proxies for the attitude of the family toward intellectual activities. These measures frequently include subscription to a daily newspaper, number of books and magazines in the home, and whether the home contains a dictionary or encyclopedia. Another family characteristic which sometimes is found in SES scales is whether the family is intact, i.e., both parents present and, if not, sometimes a distinction is made as to whether the family is broken because of death or for other reasons.

While they are not necessarily included in SES, some investigators have used number of children and position of the child in the birth order, as well as religious preference, political preference, and other variables. Since we cannot reanalyze the data in the studies we have examined, we must report the results as given in those studies. We do attempt, however, to be quite specific concerning the construction of the SES scales when such structure can be determined from the study itself.

We proceed first to discuss the aggregate amount of association that various researchers have found between SES and educational outcomes without adjustment for other factors. This discussion will not pay much attention to the composition of the SES measures. Later we will present the major findings with respect to specific measures of SES and the construction of scales from questionnaire data.

2.2.1 Composite Measures of SES

In this subsection we report some findings with respect to the association between measures of SES and various measures of educational outcomes without regard to the specific composition of the SES measures. The objective is to establish, if feasible, an amount of association that can reasonably be expected to be shown between SES and outcomes. The subsections following this one will investigate the literature with respect to the measurement of SES.

Project TALENT (14) used data from student information sheets for a national sample of about 2,900 twelfth-grade males concerning 25 items reflecting socioeconomic status. The items reflected *occupation*: father's occupation, number of people supervised, responsibilities, business and professional associations; *income*: family income and finances, rent (or value) of home, number of rooms; *cultural and education items*: books, news magazines, culture magazines, literary magazines, cultural equipment, father's education, mother's education, student has own room; *other items in home*: number of appliances, TV, radio, etc., luxury items, sports equipment, number of hand tools, number of power tools, number of cars, and year of car. These items were correlated with 14 measures of ability, and the percent of variation in each test accounted for by the socioeconomic variables is shown in Table 1.

Table 1. Squared multiple correlations between environment and various measures of ability, subsample of Project TALENT (twelfth-grade males, student is the unit of analysis)

Dependent Variable -- Abilities	Squared Multiple Correlation
Information, Part II	0.28
Information, Part I	0.27
Reading Comprehension	0.23
Math Total	0.21
English Total	0.19
Mechanical Reasoning	0.17
Abstract Reasoning	0.15
Creativity	0.14
Word Functions	0.14
Disguised Words	0.12
Visualization in Three Dimensions	0.11
Arithmetic Computation	0.10
Memory for Words	0.06
Visualization in Two Dimensions	0.06

Source: Project TALENT: One-Year Followup Studies (14, p. E-8)

It is interesting to observe that about 27-28 percent of the variation in the information-test components (corresponding roughly to accumulated knowledge) is accounted for by measures of SES and that this percentage declines as the test components become more nearly abstract measures of ability.

Bachman et al. in Youth in Transition (1) used many of the same measures of socioeconomic status as the Project TALENT socioeconomic variables in a national sample of 2,213 tenth-grade boys. About 19 percent of the variation in individual-student scores on the Quick Test of Intelligence was explained by the socioeconomic variables (1, Vol. II, p. 50). These variables were:

- Father's occupational status (coded by Duncan's scheme (11))
- Father's education (five scaled classes)
- Mother's education (five scaled classes)
- Possessions in the home (19 items)
- Number of books in the home (six classes)
- Number of rooms per person in the home

The overlap with Project TALENT items is obvious, and the results compare favorably with the results of that study.

It should be noted that in both Project TALENT and in Youth in Transition the outcome measures were ability measures rather than measures of academic performance. The title of this chapter, "Correlates of Academic Performance" might imply that measures of ability should not be used as outcome measures here. In Section 2.3.4 we consider measures of ability as background variables, but elsewhere in this report one will find some measures of ability used as measures of educational outcome.

Some of the tests in the battery used by Project TALENT (see Table 1) are highly performance oriented, and even the Quick Test used by Bachman, et al., is word-oriented so that there is a nontrivial component of accumulated knowledge in them. Therefore, in this section we have used them as proxies for achievement. Doing so has permitted us to use the particularly rich background data of Project TALENT, Youth in Transition and selected other studies. We feel that the sacrifice of logical purity is more than offset by the increased amount of data available to us.

Also the National Center for Health Statistics in its Health Examination Survey did a special study (22) of over 7,000 children in the age group 6 to 11 where some of the outcome measures reflected mental growth and behavioral development. The Wechsler Intelligence Scale for Children (WISC), the Thematic Apperception Test and the Draw-a-Person test were used. It was shown that outcomes were associated with region, with race, with income and with education of parents in a manner which compares favorably with associations of those factors with achievement tests reported in this paper.

Thorndike, in Reading Comprehension Education in Fifteen Countries (48, p. 76, 81), used father's occupation, father's education, mother's education, availability of reading matter in the home, number of siblings, and two unusual variables -- parental interest and parental help with homework (as reported by the student) -- to achieve the following percentages of explained variance on reading-comprehension tests.

	SES Alone	SES and Other Background
Ten-year-old pupils	10.9	17.5 percent
Fourteen-year-old pupils	14.4	25.8 percent
Pupils in last year of secondary school	13.0	21.0 percent

The results are for the students tested in the United States. Explained variance is again reasonably consistent with the studies reported above, although the included items are not identical and two additional age groups have been added. Note the closeness of correspondence of the three age groups to the 9, 13, and 17-year-old age groups of interest to NAEP.

Other studies in the series of International Studies in Evaluation have found similar results. Comber and Keeves (7, Chapter 9) in the science evaluations defined a "school handicap" in terms of occupation and education of father, education of mother, use of dictionary in the home, number of books in the home, and family size. When the student was the unit of analysis, the explained variance for science scores for a national U.S. sample was:

Ten-year-olds	16 percent
Fourteen-year-olds	16 percent
High school seniors	9 percent

The 9 percent for high school seniors seems unusually low. However, if sex and age are added as variables, the percentage increases to 18 percent for 10-year-olds, 22 percent for 14-year-olds, and 18 percent for seniors, results which are internally consistent and more nearly comparable with Thorndike's findings, above. The sex differential in science is known to be pronounced at the high-school-senior level, so, viewed in that light, the results are not surprising.

Percentages of explained variance increase considerably when the school is considered to be the unit of analysis. Comber and Keeves (7, Chapter 8) found the following percentages of explained variance in average science scores when SFS measures (school handicap) were averaged for each U.S. school in the sample and the average school science score was the dependent variable:

	School Handicap Only	School Handicap, Sex and Age
Ten-year-olds	66 percent	67 percent
Fourteen-year-olds	62 percent	67 percent
High school seniors	36 percent	44 percent

Note the (approximately) tripling of explained variance for 10 and 14-year-olds by using the school as the unit for analysis. Note also the relatively low percent of variation accounted for in the average science scores of high school seniors. One might hypothesize that selectivity (i.e., losses due to dropout) and the educational process have created a different explanatory system for high school seniors' science performance than for that of younger students.

We have found it necessary in many cases to use results in which the school, rather than the individual student, has been used as the unit of analysis. Since the difference in magnitude of the associations is so great, we have identified the unit of analysis as the school whenever that unit has been used.

Note that the higher associations when the school is the unit of analysis occur partly because the base of the percentage is the variation among schools means, not total variation among students. Thus, if school handicap, as defined by Comber and Keeves, accounts for 66 percent of the variance among schools, and only 25 percent of the total variance in scores is between school variance, one can approximate roughly the percent of variance in total student scores accounted for by school handicap as 0.25 x 66 = 16.5 percent. Unfortunately, the approximation is not very good and the proportion of variance between schools is usually not reported when the school is used as the unit of analysis. Mayeske (31-34) and Purves (41) are notable exceptions.

Purves (41, Chapter 6) in the international evaluation of literature scores used approximately the same background factors as Comber and Keeves. He only analyzed two age groups, 14-year-olds and high school seniors, and found the following percentages of variance in individual student scores explained by individual home background:

	Home Background Alone	Home Background, Reading Resources, Age and Sex
Fourteen-year-olds	10.2	18.5
High school seniors	7.3	16.5

Note the similarity to the results reported by Comber and Keeves and, again, the decline in association for high school seniors.

One of the best known studies of educational outcomes is the Equal Educational Opportunity Survey, the Coleman report (5). Results obtained by Coleman and his colleagues are interesting in their own right, but some of the reanalyses of his data by others are of particular relevance to this project.

Mayeske's analysis of the Coleman data (31, 32) showed that the percent of variation in a composite measure of achievement for individual students accounted for by background variables was as shown in Table 2.

Table 2. Percent of variation in composite achievement measure accounted for by background variables -- Coleman data

	Grade				
Background	12	9	6	3	1
1. SES	23	29	25	24	15
2. Family structure and stability plus (1)	23	30	27	26	15
3. Race plus (2)	32	36	37	31	22
4. Sex plus (3)	32	37	37	31	22

Source: Mayeske (31, pp. 16, 17)

These results are somewhat higher than those found in the international studies, and, for twelfth-graders, somewhat below the Project TALENT results, above.

Note also that race contributes *an additional* 5 to 10 percentage points to the percentage explained. None of the previously reported studies above have included race, and more will be said about the effect of race later.

Mayeske used factor-analytic methods for creating an index of achievement and then used criterion-scale analysis to create scales for questionnaire items that reflected SES. That is, since 12th-grade students who reported "manager" as their father's occupation had an average standardized achievement index of 52.771, that value was assigned as the numeric code for "manager." This technique tended to maximize the associations between the components of the SES measure and the index of achievement. The results from the factor analyses were then used in subsequent regression analyses. It is possible that the assignment of criterion scales to questionnaire items and the compositing by factor-analytic techniques produced the apparent greater proportion of explained variance than was found by the methods of compositing used in the international assessments. The same set of student scores was used by Mayeske in forming his achievement index as in forming his criterion scales. This must have added something to predictability, even though the sample was large.

Hilton, et al. (23), in a compilation of data from the Educational Testing Service "Academic Growth Study," used analysis of variance techniques in order to partition the total sum of squares of deviations of student scores from their overall means into "main effects,' "first order interactions," "second order interactions," etc., in the analytical pattern often used in experimental designs. A wide array of tests was used and those tests that were keyed to academic performance showed high significance of the father's education factor which was used as the measure of SES. Also, the interactions of that factor with many of the other background factors tended to be significant.

The forthcoming final reports on the first wave of the "National Longitudinal Study of the High School Class of 1972" (25) will also show associations with SES that are comparable with the principal results cited in this section. As of the date of preparation of this study the National Longitudinal Studies have not been released to the public so no specific data have been quoted in this report. However, the researcher interested in background factors should be aware of this important forthcoming series of reports.

We could cite other findings in the literature to summarize the level of association between measures of SES and measures of academic outcome. However, we feel that the importance of SES and the amount of association to be expected in the explanation of differences in academic outcomes has been demonstrated conclusively by the references cited. We move now to a consideration of the ways in which SES is measured.

2.2.2 Components of SES

In considering measures of components of SES, a reality to face is that one often can't measure the thing he really wants to measure but must settle for measuring one or more proxies. For example, he might like to know family income, but even carefully trained interviewers using multiple questions that, for example, distinguish between wages and take-home pay and identify other than wage income are unable to ascertain family income accurately. It certainly is expecting too much of younger students to have them report family income even if such questions were permitted in all of the states. The result is that instead of asking students to report income one asks them to report such things as occupation of parents, automobiles and items in the home that reflect standard of living. These items are correlated with income and may serve as reasonable proxies for income.

We now consider the following major components of SES: occupation of parents, items in the home, and education of parents.

2.2.2.1 *Occupation of Parents*. One of the most common surrogates for income is occupation of father, or occupation of the mother if there is no father in the home. With the higher percentage of women working in recent years it seems surprising that more attention has not been given to whether the mother is working (full- or part-time) and to her occupation. The fact that the mother is working may have an impact on certain educational outcomes quite aside from the impact of a greater family income.

The International Studies in Education have used ten occupational categories. The categories were designed by C. A. Anderson for the mathematics evaluation (26, Vol. I, Chapter 8) and have set the pattern for subsequent evaluation. The complete classification is given in the referenced report, but the category names are given here:

Group 1 - Higher professional and technical

Group 2 - Administrators, executives, and working proprietors, large- and medium-scale

Group 3 - Subprofessional and technical

Group 4 - Small working proprietors (*other than in agriculture, forestry, or fishing*)

Group 5 - Proprietors and managers in agriculture, forestry, and fishing

Group 6 - Clerical and sales workers (lower levels of white-collar work)

Group 7 - Manual workers, skilled and semiskilled

Group 8 - Laborers (hired) in agriculture, forestry

Group 9 - Unskilled manual workers (excluding agriculture, forestry, fishing)

Group 0 - Unclassified; no answer.

Students were asked to "describe your father's occupation carefully" and one of the above codes was assigned as part of the data processing.

At a higher level of aggregation one can combine (1 and 2), (3, 4, and 6), (7 and 9) and (5 and 8). B. S. Bloom et al., (*International Study of Achievement in Mathematics: A Comparison of Twelve Countries*, Volume II, Chapter 5, "Social Factors in Education") (26) showed that achievement in mathematics in the United States was approximately linearly related to these aggregations where the aggregates are listed in reverse order of "size," above, and are assumed to form a scale of equal units.

Peaker (*International Study of Achievement in Mathematics: A Comparison of Twelve Countries*, Volume II, Chapter 6, "A Regression Analysis") (26), in doing a regression analysis of mathematics achievement on occupations, used a second dimension, viz., whether the occupation was scientific or otherwise. He showed that for the United States the regression coefficient for the second dimension had small value and was negative, indicating that the nine-category scale contained effectively all of the information in the scientific/nonscientific classification as well. Although the text is not specific on the point it is believed from a remark in an earlier chapter that for the purpose of regression analysis scale values of one to nine were assigned to the categories in reverse order of the above listing.

Because of difficulties encountered in the mathematics evaluation in constructing a scale of occupations that was valid cross-nationally, the international literature evaluation (Purves, (41)) requested each country to construct its own scale. When these scales had obvious shortcomings the technique of criterion scaling as described above in connection with the Mayeske report was used.

In the international reading-comprehension evaluation (48) each country developed its own categories. In the U. S. they were as follows:

Code Number	Category
9	Professional, technical and kindred workers
8	Managers, officials and proprietors, including farm owners and managers
7	White-collar workers
6	Skilled manual workers
5	Semiskilled manual workers
4	Farm workers, fishery, forestry and kindred groups
3	Domestic and personal service workers
2	Laborers
1	Unclassifiable
0	Unknown

---

[1] See A. E. Beaton's paper in Mayeske et. al., A Study of Our Nation's Schools (32, Appendix).

Coding was done centrally on the basis of the student's report of the specific job that his father held. The method of converting to a numerical score is described by Thorndike as follows:

> In order to translate the coding of Father's occupation into a score that would be an effective predictor of reading achievement, a regression analysis was carried out of the several categories of the score, assigning each individual a code of 1 for the category in which his father's occupation fell and a code of zero for all other occupations. So, in the United States, the nine categories of occupation were treated as nine dummy variables, and that weighting was determined for this set of nine variables that would maximize the correlation of the set with Reading Comprehension score. Thus, Father's occupation was empirically scaled in such a way as to maximize its predictive effectiveness.

Scaling of father's occupation was done for the 14-year-old population and that scale was also used for the 10-year-old population and for high school seniors.

Comber and Keeves (7), in the international science evaluations, used the same technique of criterion scaling as Purves, above. Thus, we see that the framework for the international studies is the nine-category classification of occupations based upon descriptive reporting by students and centralized coding and (for recent studies) assignment of scale values by the method of criterion scaling.

Comber and Keeves (7, Chapter 9) report the following percentages of variance in individual student scores in science accounted for by father's occupation (U.S. data):

Ten-year-olds	15.2
Fourteen-year-olds	9.6
High school seniors	4.4

Again, we note the decline with increasing age that has been observed earlier. These percentages compare with 16, 16 and 9 percent, for the three age groups, respectively, that Comber and Keeves attributed to occupation and education of father, education of mother, use of dictionary in the home, number of books in the home and family size. Thus, for ten-year-olds, occupation of father accounts for substantially all of the variation of the larger set, and for high-school seniors, to about half of the variation of the larger set of variables.

The Coleman studies (5) elicited information from school personnel concerning the occupational status of the student's father for first- and third-grade students and from the students themselves for sixth, ninth, and twelfth grades. In both cases a precoded check response was sought rather than a verbal description which had to be coded subsequently. The categories for the student response were as follows:

18. What work does your father do? You probably will not find his exact job listed, but check the one that comes closest. If he is now out of work or if he's retired, mark the one that he usually did. Mark only his main job if he works or more than one.

    (A) Technical -- such as draftsman, surveyor, medical or dental technician, etc.

    (B) Official -- such as manufacturer, officer in a large company, banker, government official or inspector, etc.

    (C) Manager -- such as sales manager, store manager, office manager, factory supervisor, etc.

    Proprietor or owner -- such as owner of a small business, wholesaler, retailer, contractor, restaurant owner, etc.

    (D) Semiskilled worker -- such as factory machine operator, bus or cab driver, meat cutter, etc.

    Clerical worker -- such as bank teller, bookkeeper, sales clerk, office clerk, mail carrier, messenger, etc.

    Service worker -- such as barber, waiter, etc.

    Protective worker -- such as policeman, detective, sheriff, fireman, etc.

(E) Salesman -- such as real estate or insurance salesman, factory representative, etc.

(F) Farm or ranch manager or owner

(G) Farm worker on one or more than one farm

(H) Workman or laborer -- such as factory or mine worker, fisherman, filling station attendant, longshoreman, etc.

(I) Professional -- such as accountant, artist, clergyman, dentist, doctor, engineer, lawyer, librarian, scientist, college professor, social worker, etc.

(J) Skilled worker or foreman -- such as baker, carpenter, electrician, enlisted man in the armed forces, mechanic, plumber, plasterer, tailor, foreman in a factory or mine, etc.

(K) Don't know

Suitable modifications in wording were made for the school response for first and third grades.

As described earlier, Mayeske (31, 32) used criterion scaling to assign values to categories A through K above, different values being assigned for each grade level. His criterion scale values are shown in Table 3. He noted that twelfth-graders who did not know their father's occupation had a lower achievement index than first, third or sixth-graders. This assignment of separate criterion scales to each grade may have been one of the procedures that gave Mayeske a higher explained variance than was found by the researchers in the international evaluation studies. The effect of that procedure is confounded with the fact that the Coleman survey used precoded categories and the international studies used descriptive language which was coded into categories in the data-processing phase.

Percents of variance explained shown in the last line of Table 3 were computed from tabulations of average criterion scores and frequency tabulations provided us by Dr. Mayeske (34). The lower explained percentage for third grade may reflect the fact that

Table 3. Scale values assigned by Mayeske to occupation of father -- Coleman data

Father's Occupation	Means for the Different Grade Levels			
	Third	Sixth	Ninth	Twelfth
A. Technical	54.4	48.7	52.7	52.4
B. Official	54.5	52.8	52.3	52.7
C. Manager	54.6	53.6	53.5	52.8
D. Semiskilled	50.1	50.0	50.1	49.5
E. Salesman	53.9	54.1	53.9	53.6
F. Farm or ranch manager or owner	52.9	50.2	50.4	50.7
G. Farm worker	45.7	45.5	43.3	42.5
H. Workman or laborer	45.9	49.6	48.7	47.2
I. Professional	56.8	55.3	56.6	56.0
J. Skilled worker or foreman	50.8	51.4	51.0	50.6
K. Don't know	45.7	44.1	43.1	41.8
Nonresponse	48.8	45.0	42.6	42.3
Total	50.0	50.0	50.0	50.0
Percent of variance explained	10.4	12.0	15.0	12.2

occupation of father was obtained from teachers for third graders and from the students themselves in other grades. The drop off in explained variance for twelfth graders will be noted frequently in this report. It has been suggested that high school dropouts tend to make the senior class more homogeneous than other grades.

Bachman, in *Youth in Transition* (51, Volume II), used Duncan's scale (Reiss (42)). Categories were assigned by coders based upon verbal description of father's occupation. Since the use of Duncan's scale is frequently encountered in the literature, it seems desirable to discuss it briefly here.

Duncan used two variables in the construction of his index, both available from the 1950 Census. The first was percent of the persons in the occupational category receiving more than $3,500 in income, and the second was percent with at least a high school education. Some adjustment was made in income distribution to standardize to a given age. The index is found by weighting the first variable by 0.59, the second by 0.55 and subtracting 6.0, these constants being obtained by regression methods using the NORC prestige scale as a criterion (Reiss, 42). The scale value was produced for each occupation in the detailed occupational classification of the Census. The complete scale is given in Appendix B of Reiss (42).

Flanagan et. al., in Project TALENT (13), asked several questions related to occupation:

131. Does your father work for pay on more than one job?

132. Does your father direct or supervise the work of other people?

133. As far as you know, which one of the following best describes your father's responsibility for money and property on his job? (Four categories plus "don't know")

134. Has your mother worked for pay at any time in the last three years?

135. How long has your mother been working for pay?

How active has your father been in any one or more of the following organizations?

141. Labor-union or trade-union activities?

142. Business or professional association? (Questions 151 and 152 were parallel to 141 and 142, except with reference to mother's activities)

In addition, two occupational questions were asked. Question 206 identified the father's occupation in one of 17 categories as follows:

A. Farm or ranch owner and/or manager

B. Farm or ranch foreman

C. Farm or ranch worker

D. Workman or laborer -- such as factory or mine worker, fisherman, filling station attendant, longshoreman, etc.

E. Private household worker -- such as servant, butler, etc.

F. Protective worker -- such as a policeman, detective, sheriff, fireman

G. Service worker -- such as barber, beautician, waiter, etc.

H. Semiskilled worker -- such as factory machine operator, bus or cab driver, meat cutter, etc.

I. Skilled worker or foreman -- such as baker, carpenter, electrician, enlisted man in the armed forces, mechanic, plumber, plasterer, tailor, foreman in a factory or mine (but not on a farm), etc.

J. Clerical worker -- such as bank teller, bookkeeper, sales clerk, office clerk, mail carrier, messenger, etc.

K. Salesman -- such as real estate or insurance salesman, factory representative, etc.

L. Manager -- such as sales manager, store manager, office manager, business manager, factory supervisor, etc.

M. Official -- such as manufacturer, officer in a large company, banker, government official or inspector, etc.

N. Proprietor or owner -- such as owner of a small business, wholesaler, retailer, contractor, restaurant owner, etc.

O. Professional -- such as actor, accountant, artist, clergyman, dentist, engineer, lawyer, librarian, scientist, etc.

P. Technical -- such as draftsman, surveyor, medical or dental technician, etc.

Q. I don't know.

If the respondent answered O or Q, above, he was given, in question 207, a list of 35 professions to check. The same questions were repeated for mothers if they were working for pay.

In Appendix E to The American High School Student (13) Flanagan reported number of students of fathers in each of the above occupational classes who fell in each decile of his ability measure. From these data it is possible to estimate the percent of variance in ability (which we use as an approximation to educational performance) accounted for by father's occupation. The computations were made by us only for ninth grade girls and are shown in Table 4.

Table 4. Relationship between father's occupation and measures of ability for ninth grade girls -- Project TALENT data

Occupational Category	Number of Observations	Average Percentile Rank
A. Farm or ranch owner/manager	963	46.88
B. Farm or ranch foreman	136	24.08
C. Farm or ranch worker	436	31.19
D. Workman or laborer	2,128	47.80
E. Private household worker	54	17.59
F. Protective worker	115	47.95
G. Service Worker	127	45.70
H. Semiskilled worker	654	51.86
I. Skilled worker or foreman	1,631	56.86
J. Clerical worker	212	59.95
K. Salesman	321	65.84
L. Manager	468	62.39
M. Official	135	57.96
N. Proprietor	561	65.81
O. Professional	382	66.47
P. Technical	180	67.06
Q. Don't know	1,579	35.08
X. Nonresponse	766	31.12

Estimated percent of variance accounted for = 16.0

In the one-year followup study (14), items 132, 133, 142, 206, and 207 were used in the measure of SES, along with 21 other items. Simple correlations with the two parts of the "Information" test showed that the responsibilites response, question 133, had a correlation of 0.14 with both parts of the information test and question 206 a correlation of 0.26 with one part and 0.24 with the second part of the information test. Other simple correlations were less than 0.10 and were not reported. Note that the above figures are correlation coefficients and not percents of explained variance (squares of correlation coefficients). Thus, father's occupation alone accounts for six of seven percentage points of the 27 or 28 percent of the individual variance in the Information parts of the tests accounted for by the full set of environmental variables (see Table 1). Only father's occupation was included in the final Socioeconomic Environment Index for the one-year followup study with the following scaling:

Scale Score	Item 206 Response
1	C. Farm or ranch worker D. Workman or laborer E. Private household worker
2	G. Service worker H. Semiskilled worker
3	B. Farm or ranch foreman F. Protective worker I. Skilled worker or foreman J. Clerical worker
4	A. Farm or ranch owner K. Salesman L. Manager N. Proprietor or owner P. Technical
5	M. Official O. Professional

This scale apparently was used in computing the correlation cited above.

From the cited references it appears that occupation of father is a significant component of SES. Alone it can account for from (say) 5 to 15 percent of variation in academic performance measures.

2.2.2.2 *Items in the Home*. Since it is difficult, and often impossible, to obtain information on family income directly, a commonly used surrogate is items in the home such as appliances, television, automobile, and so on. The possession of a number of such items is presumed to be correlated with family income. In addition, the possession of some items is presumed to reflect an attitude toward learning on the part of the parents. An encyclopedia, a dictionary, and a subscription to a daily newspaper qualify as items of this kind. Some items are quite specifically related to the subject matter that is the focus of NAEP exercises. Examples are subscriptions to science, literary, music and art magazines, the possession of musical instruments, paintings, classical-music recordings, and so on. Most items, however, are assumed to be correlated with academic performance in a more general way.

At the present time, NAEP is requesting information on the following items in the home:

Daily newspaper
Magazines (regularly)
More than 25 books
Encyclopedia

We understand that NAEP chose these items on the basis of a study which collected data on a larger number of such items and which investigated association with other measures of SES.

Examination of some of the principal research papers relating educational outcomes to background impresses one with the variety and complexity of items for which data have been collected and the paucity of careful analysis of the relationships. Even though a classification of household items is artificial, it is helpful to group items according to whether they reflect primarily income, general home educational environment, or special home educational environment. Obviously, items that are present in nearly all homes, or in very few homes, cannot serve as useful indications of SES level, but may be helpful in identifying small subgroups with special interests. Some of the items requested in principal studies are the following (5, 1, 13):

Income Indicators

Television set
Telephone
Refrigerator
Automobile
Vacuum cleaner
Rooms in home
Separate room for student
Bicycle
Automatic washer
Clothes dryer
Dishwasher
Food freezer
Sterling silver
Wall-to-wall carpeting
Ceiling-to-floor draperies
Various identified sports equipment
Power tools

General Home Educational Environment

Dictionary
Encyclopedia
Daily newspaper
Number of books in home
Magazine subscriptions
Map or globe of the world
Typewriter

Special Home Educational Environment

Record player, hi-fi or stereo
Bible
Camera
Dog or cat
Fish in a tank
Pair of binoculars
Number of phonograph records
Specifically named magazines
Identified kinds of magazines (10 kinds)
Paintings

Tapestries
Musical instruments
Art equipment
Photographic development equipment
Hand tools

Clearly, the classification of many items is arbitrary and the list is by no means exhaustive.

Armor in his reanalysis of the Coleman data reported correlations of sixth-grade verbal achievement with the household items index formed by giving one point each for possession of television, telephone, record player or hi-fi, refrigerator, dictionary, encyclopedia, automobile, vacuum cleaner and daily newspaper. The data are summarized in Table 5.

Table 5. Correlation of school-average sixth-grade achievement scores with school means of the household items index -- Coleman data

Class	r	r^2
Black achievement in black schools		
Metropolitan - North	0.59	0.35
Metropolitan - South	0.54	0.29
Nonmetropolitan - South	0.39	0.15
White achievement in white schools		
Metropolitan - North	0.71	0.51
Metropolitan - South	0.66	0.44
Nonmetropolitan - North	0.38	0.14
Nonmetropolitan - South	0.72	0.52

Source: Armor (37, p. 213)

Note that the school rather than the student is the unit of analysis in the computation of the correlations in Table 5. We are, of course, most interested in the correlations among individual characteristics and outcomes, and as we have seen, correlations such as the above for school averages account for a much smaller fraction of the individual outcomes than is indicated by the squared correlation for schools.

Mayeske (31, 32) in analyzing the Coleman data assigned a scale to items in the home by use of criterion scaling (Beaton, 32, Appendix I) and by reduction in number of items by factor-analytic methods. Separate scales were constructed for first, third, sixth, ninth, and twelfth grades. The scale values are given in Table 6 for third, sixth, and twelfth-graders, grades which roughly coincide with age levels 1, 2, and 3 of the National Assessment. The scale values have been rounded to one decimal point from three for convenience.

These scale values are averages of the criterion (achievement composite) standardized to a mean of 50 and standard deviation of 10. Thus, an item has potential to discriminate between high and low outcome if there are large differences in the scale values for the various response categories in Table 6. All of the items seem to be related to outcomes. Among the appliances, the one that seems to have greatest discriminating capability is presence or absence of a refrigerator. However, the number of "no-refrigerator" responses is quite small. The amount of consistency among the items leads one to believe that the aggregate of a simple scale formed by giving a score of 1 for presence and 0 for absence (as was done by Armor) might be reasonably effective. The scale values assigned to number of magazines and number of books are approximately linear on the classes employed if one combines "7 or more" magazines with "5 or 6." It should be noted, however, that there is nothing in the criterion-scaling procedure which permits one to assign greater weight to one item (say, dictionary) than to another item (say, encyclopedia) if both are to be used in the composite. Such differential weights can be determined in a subsequent analysis by regression analysis, factor analysis or other methods to aid in judgments.

The percent of houses having an item is so important to the amount of variance explained by that item that it has been analyzed for twelfth graders in Table 7. Thus we see that, as mentioned above, lack of a refrigerator in the home marks the student from that home as one with an educational handicap, but less than two percent of homes reported no refrigerator. As a result, possession of a refrigerator only accounts for 1.5 percent of the variance in the criterion.

Table 6. Scale values used by Mayeske for items in the home -- Coleman data

Item	Response	Grade 3	Grade 6	Grade 12
Appliances				
Television	Yes	50.4	50.6	50.2
	No	44.5	42.6	45.2
	N.R.	42.1	37.7	43.2
Telephone	Yes	51.4	51.6	50.9
	No	45.2	44.1	43.2
	N.R.	42.5	37.1	42.9
Record player, hi-fi, or stereo	Yes	50.8	51.1	50.4
	No	47.4	45.4	47.3
	N.R.	43.6	36.9	43.3
Refrigerator	Yes	50.3	50.9	50.2
	No	43.5	38.3	39.5
	N.R.	42.5	36.7	43.5
Automobile	Yes	51.6	51.2	50.6
	No	44.6	42.3	42.5
	N.R.	44.6	36.1	42.8
Vacuum cleaner	Yes	51.6	52.0	
	No	45.2	43.7	
	N.R.	44.4	36.5	
Reading Materials in Home				
Dictionary	Yes	51.3	51.2	50.3
	No	45.6	42.9	41.7
	D.K.	48.6	39.7	-
	N.R.	42.9	36.5	43.7
Encyclopedia	Yes	51.9	52.2	51.0
	No	46.8	46.2	46.1
	D.K.	49.0	40.7	-
	N.R.	44.3	36.9	41.6
Daily newspaper	Yes	51.1	51.5	50.8
	No	48.2	46.3	45.8
	N.R.	44.2	39.2	44.4
Number of magazines	None			45.9
	1 or 2			48.5
	3 or 4			50.9
	5 or 6			53.6
	7 or more			52.7
	N.R.			42.5
Number of books	0-9			44.5
	10-24			45.9
	25-99			49.7
	100-249			52.6
	250 or more			54.4
	N.R.			42.8

Source: Mayeske (32, Appendix III)

Table 7. Percent of students reporting items in the home and percent of variance explained for twelfth-graders -- Coleman data

	Percent answering			Percent of Explained Variance
	Yes	No	Nonresponse	
Appliances				
Television	96.8	2.5	0.7	0.9
Telephone	88.2	11.0	0.8	6.3
Record player, hi-fi or stereo	88.4	10.8	0.8	1.3
Refrigerator	98.1	1.1	0.7	1.5
Automobile	92.8	6.4	0.8	4.4
Vacuum cleaner	85.0	14.2	0.8	9.1
Reading materials in home				
Dictionary	96.9	2.3	0.8	2.0
Encyclopedia	80.4	18.8	0.9	4.4
Daily newspaper	85.1	13.9	1.1	3.3
Number of magazines	*	*	*	6.2
Number of books	**	**	**	10.1

*		**	
None	12.6	0-9	5.8
1 or 2	30.5	10-24	18.8
3 or 4	33.0	25-99	36.6
5 or 6	14.3	100-249	21.6
7 or more	8.6	250 or more	16.1
Nonresponse	1.0	Nonresponse	1.1

Presence or absence of a vacuum cleaner, on the other hand, accounts for 9.1 percent of the variance in the criterion. Number of books in the home emerges as the best single predictor for twelfth-graders, accounting for over 10 percent of the variance.

Items in the home undoubtedly have different predictability for various population subgroups and that difference is not exposed by the above analysis.

Marshall Smith (37, Chapter 6) separated the Coleman data on items in the home into (1) reading material and (2) other items. Although the text is not specific, he appears to have used simple counts of items as Armor did. He reported simple correlations between these variables and individual verbal achievement, as shown in Table 8.

Table 8. Simple correlations of reading material and items in the home with verbal achievement -- Coleman data

Group	Reading Material	Items in Home
Sixth Grade		
Northern Black	0.21	0.23
Northern White	0.23	0.22
Ninth Grade		
Northern Black	0.16	0.15
Northern White	0.27	0.17
Twelfth Grade		
Northern Black	0.15	0.13
Northern White	0.29	0.13

Source: Smith (37, Chapter 6, Appendix)

The multiple correlations, using both reading material and items in home, are not given. It should be noted that correlations across students in all schools would be expected to be somewhat higher than those shown in Table 8.

Project TALENT collected data on many items in the home with particular emphasis on kinds of magazines regularly taken. Unfortunately, there is limited published data on the associations between possession of specific items and outcomes. The thrust of Project TALENT was to identify ability, so measures of ability must be used in place of academic performance. However, many of the tests are performance-oriented. A composite measure of academic aptitude was used as a criterion variable in an analysis which permits one to draw an association with number of books in the home (13, Appendix E). Table 9 gives median number of books in the home for deciles of the academic-aptitude criterion.

Table 9. Median books in the home (excluding non-response) for deciles of the Project TALENT academic-aptitude criterion

	Median Books in Home			
Aptitude Decile	9th Grade Boys	9th Grade Girls	12th Grade Boys	12th Grade Girls
1	51	25	62	47
2	39	44	51	49
3	53	44	58	40
4	53	54	63	59
5	64	58	65	60
6	65	58	65	67
7	72	67	73	71
8	79	72	83	77
9	93	73	95	89
10	118	102	123	147

Source: Project TALENT (13, Appendix E)

Also Cooley, Dalcanton, McMillen et. al. (14, Appendix E) reported simple correlations between selected items in the home and various measures of ability. The two "information" components of the ability tests are probably closer to measures of academic performance than are the other ability tests. Correlations reported by Cooley et. al. are given in Table 10. Most of the questions were asked in terms of how many of the given items were in the home. Most questions provided for answers of none, one, two, three, four, five or more. Categories for number of books were 0-10, 11-25, 26-100, 101-250, 251-500 and 501 or more.

Table 10. Selected correlations between household items and two components of the information test, Project TALENT

Items	Correlation Coefficients	
	Information I	Information II
Books in home	0.28	0.29
News magazines	*	*
Culture magazines	*	*
Literary magazines	0.14	0.15
Appliances	0.24	0.26
TV, radio, telephone, etc.	0.33	0.35
Luxury items	*	0.10
Cultural equipment	0.17	0.16
Sports equipment	0.17	0.18
Own room and desk	0.30	0.30
Hand tools	0.17	0.14
Power tools	*	*
Cars owned	*	*
Year of car	0.12	0.14

* Less than 0.10.

Source: Project TALENT (14, Appendix E)

When the correlations in Table 10 are squared one can see that the percentage of variation explained by any one of the items does not exceed 12 percent. However, the four relatively most important associations with the information test are, in order of importance:

T.V., radio, telephone, etc.
Own room and desk
Number of books in home
Number of appliances

Intercorrelations among these four items were reported as shown in Table 11.

Note that number of books is poorly correlated with the other variables and hence would contribute substantially to variance explained by the others. The authors

Table 11. Correlations among items in home -- Project TALENT data

Items	1	2	3	4
1. TV, radio, telephone, etc.	1.00	0.32	*	0.46
2. Own room and desk	0.32	1.00	*	0.30
3. Number of books in home	*	*	1.00	*
4. Number of appliances	0.46	0.30	*	1.00

* Less than 0.30.

Source: Project TALENT (14, Appendix E)

finally selected (1) books in home, (2) number of appliances, (3) TV, radio, telephone, etc., and (4) own room (the four identified above) to be included in their socioeconomic environment index. The primary selection criterion was that "...the items selected had to be answerable by most of the students and had to measure aspects of the environment closely related to student ability." The final socioeconomic environment (SEE) index contained five other items as well (value of home, family income, father's occupation, father's education and mother's education) and was formed by adding together the standard scores for the nine items (i.e., standardized to mean 100 and standard deviation 10). The weightings that would have resulted from a regression analysis are unknown and cannot be computed from the published data.

Bachman et. al. in Youth in Transition recorded presence or absence of 19 items in the home with the percent reporting them as shown in Table 12. A six-point scale was constructed as follows:

9 or fewer items	1
10 or 11 items	2
12 or 13 items	3
14 or 15 items	4
16 or 17 items	5
18 or 19 items	6

Table 12. Percent of students' homes containing items reflecting SES -- Bachman data

Item	Percent Having	Item	Percent Having
Radio	97	Typewriter	66
Telephone	90	Dog or cat	67
Television	96	Fish in a tank	20
Bicycle	84	Daily newspaper	79
Phonograph	87	Magazine subscription	79
Dictionary	96	Pair of binoculars	49
Set of encyclopedias	81	More than 10 phonograph records	88
30 other books or more	86	Map or globe of the world	81
Family car	92		
Camera	92		

Source: Youth in Transition (1), Volume II, p. 12)

In addition, number of books in the house was obtained and scaled as follows:

None, or very few (0-10)	1
A few (11-25)	2
One bookcase full (26-100)	3
Two bookcases full (101-250)	4
Three or four bookcases full (251-500)	5
A room full -- a library (501 or more)	6

Simple correlation coefficients are reported as follows:

	1	2	3
1. Quick Test	1.00	0.34	0.28
2. Items in home	0.34	1.00	0.43
3. Books in home	0.28	0.43	1.00

These results are reasonably in line with the results reported by Flanagan et. al., above, for Project TALENT and summarized in Table 10. Note again that the Quick Test is an ability test, rather than a performance test, but it clearly is related to accumulated knowledge.

Some simple manipulation of the above figures shows that the following percentages of variation in Quick Test scores can be accounted for:

By items in the home alone	12 percent
By books in the home alone	8 percent
By items in the home and books in the home together	14 percent

These results appear to summarize reasonably well what can be expected from using items in the home as an indicator of SES.

2.2.2.3 *Education of Parents*. Education of father or mother or both has frequently been used in measures of SES. Education of parents is correlated with occupation and with income, but apparently contributes something additional to predictability of academic outcomes.

The Coleman survey (5) asked for education of each parent in terms of the following categories:

(A) None, or some grade school
(B) Completed grade school
(C) Some high school, but did not graduate
(D) Graduated from high school
(E) Vocational or business school after high school
(F) Some college, but less than four years
(G) Graduated from a four-year college
(H) Attended graduate or professional school
(I) Don't know

The survey administrator completed the data for first and third grades, and there was student response for sixth, ninth and twelfth grades.

Mayeske (31, 32) used the method of criterion scaling to assign scale values to the above categories with the results shown in Table 13 for third, sixth, and twelfth grades. Again, these values are scaled to an overall mean of 50 and standard deviation of 10. The scales for fathers and mothers show great similarities in pattern. Education of husbands and wives is, of course, correlated.

Table 13. Scale values assigned by Mayeske to parental-education categories -- Coleman data

	Grade		
Educational Level	3	6	12
Father's Education			
None, or some grade school	43.2	43.9	45.2
Completed grade school	47.3	46.4	48.2
Some high school, but did not graduate	48.2	49.4	49.0
Graduated from high school	52.1	52.4	51.7
Vocational or business school after high school	54.2	50.5	53.4
Some college, but less than four years	54.8	54.1	54.3
Graduated from a four-year college	56.2	55.1	55.1
Attended graduate or professional school	58.2	56.0	57.4
Don't know	48.3	48.7	43.9
Nonresponse	48.6	44.6	43.1
Percent of variance explained	11.2	9.8	12.5
Mother's Education			
None, or some grade school	42.6	43.3	44.5
Completed grade school	46.4	45.2	46.9
Some high school, but did not graduate	47.7	48.5	47.6
Graduated from high school	52.2	52.8	51.6
Vocational or business school after high school	55.8	52.3	54.8
Some college, but less than four years	54.8	53.7	54.4
Graduated from a four-year college	56.5	54.2	55.4
Attended graduate or professional school	55.8	52.5	55.3
Don't know	48.6	48.5	43.9
Nonresponse	48.9	44.4	43.9
Percent of variance explained	10.1	9.6	11.9

Source: Mayeske (32, Appendix III)

Bachman in Youth in Transition used the following categories of parental education:

Less than high school
Some high school
Completed high school
Some college
Completed college

A linear five-point scale was applied to the above five categories. Some key correlations were reported by Bachman as follows:

	1	2	3
1. Reading achievement	1.00	0.28	0.25
2. Father's education	0.28	1.00	0.51
3. Mother's education	0.25	0.51	1.00

Father's education alone accounts for 8 percent of the variation in reading achievement, mother's education for 6 percent, and the two together 10 percent.

Project TALENT (13) used the same education categories as Coleman, above, except that master's degree and doctorate or professional degree were added as specific categories. Although the text is not specific (14, p. E-10) it is believed that a simple linear scale was used in the analysis. It was found that both father's education and mother's education correlated about 0.27 with the information components of the TALENT ability tests and 0.45 with each other. Note the similarity to the Bachman results cited above.

Husén in the International Study of Achievement in Mathematics (26) asked, "For how many years did your father (mother) receive full-time education?" Responses were recorded in number of years and apparently these values formed the scale for analysis. This set the pattern for subsequent international studies. Husén showed a correlation between mathematics score for 13-year-olds and education of father (U.S. results) of 0.30 and 0.28 with education of mother. Again, note the consistency with the results of Flanagan and Bachman.

Education of father (or of both parents) was included as a background variable in all of the international studies as part of the measure of SES, but the unique contribution to the SES measure was generally not computed.

The results cited lead one to believe that education of both parents may account for around 10 percent of individual variation in academic test outcomes and that education of father alone might account for something like three-fourths of that.

2.3 Personal Characteristics of Students

In this section we discuss the associations between academic achievement and those characteristics of the student that are immutable such as race, sex, age, family structure and student ability, although, except for race and sex, immutability is a short-range concept.

2.3.1 Ethnic Group Membership

We suggest that race not be included as a component of SES, although Mayeske (31) showed that race contributed an additional 5 to 10 percent to explained variation after SES and family structure and stability had been accounted for. Armor (37, p. 217) has shown that percent black in the school has strong predictability as a school variable.

> It is the largest of the community input factor coefficients. Thus, percent black explains more variation in achievement even when four other socioeconomic factors are controlled. This tends to confirm the notion that black families (or communities), on the average, may be more disadvantaged vis-a-vis whites than can be determined by objective social-class measures alone.

This view seems to be widely held and forms the basis for much of the literature on what is meant by equal educational opportunity. (See, for example, Hauser, (21, pp. 13ff)).

Coleman (5, Section 3.2) has shown how important race (or ethnic group) is to the amount of variance explained by background factors. His results appear to show that background factors are differentially important for various minority groups, which suggests that adjustment for background should be done separately for the principal minority groups (and, possibly, for geographic region).

We have assumed that there are compelling reasons why race of student should be recorded and used in the analysis and therefore have not made any comprehensive search of the literature to justify the inclusion of race as a background characteristic. Instead, we have explored methods of recording race or ethnic group.

In the Coleman survey, race and ethnic group were determined by self-response (except in grade 1) to three questions as follows (for sixth-graders):

4. Which of the following best describes you?

   (A) Negro
   (B) White
   (C) American Indian
   (D) Oriental
   (E) Other

5. Are you Puerto Rican?

   (A) Yes
   (B) No

6. Are you Mexican-American?

   (A) Yes
   (B) No.

In addition, question 3 identifies children born in Puerto Rico and in Mexico as a part of a more general question concerning place of birth. Responses to the questions permit one to separate two important components of the Spanish-origin populations.

Youth in Transition (1, Volume I) required the interviewer to record after the interview his opinion as to whether the respondent was white, Negro or other (to be specified by interviewer). Recording of race by an observer is an error-prone procedure. Also, performance is so highly variable for "other" that it would seem advisable for NAEP to specify greater detail. Finally, the lack of any special treatment for Spanish-origin populations seems inadvisable for NAEP.

There might be some merit in recording race in the same manner as was done for the 1970 Census of Population so that census data might be used for estimation and benchmarking purposes. In the complete enumeration the following categories were given:

White
Negro or black
Indian (American)
Japanese
Chinese
Filipino
Hawaiian
Korean
Other (to be specified)

The detail may be too great for NAEP purposes, but might provide a classification that could be collapsed.

In the five percent sample of the 1970 Census a question concerning origin or descent was asked with the following classes:

Mexican
Puerto Rican
Cuban
Central or South American
Other Spanish
None of these

Also, in the 15 percent sample a question concerning language, other than English, spoken in the home was asked.

Finally, for some tabulations for selected areas, names were matched against a master file of Spanish surnames. If they matched, the record was used in the tabulation of characteristics of Spanish-surnamed persons. This procedure alone does not identify two components of the Spanish-origin population that are of interest, namely the Mexican-American and Puerto Rican populations.

In Year 01 NAEP obtained information on race by having the exercise administrator note race as "black," "nonblack," or "other nonwhite" on the assessment package as it was turned in. Science results were reported by these categories in Report 7 -- both balanced and unadjusted results. Race is still being recorded in this manner.

There is little doubt that interest in particular ethnic or racial subgroups is so great that NAEP will continue to record such characteristics. It seems advisable to record some additional detail for possible future analyses, however.

2.3.2 Age and Sex

NAEP is a study of 9, 13, and 17-year-olds, regardless of the grades in which they are enrolled, and of young adults. Thus, age varies only within a narrow range, but the grade in which enrolled is a variable. The grade in which enrolled is largely a function of school policy, but also reflects academic progress of the child, parental attitude, student health, and other factors. One would expect that the greater amount of educational experience of (say) nine-year-old students in the fourth grade than of nine-year-old students in the third grade would have an impact on performance. It seems apparent, then, that NAEP should record the grade in which enrolled.

The age at which the student began the first grade might also be a variable with significant explanatory power. (See, for example, effect of age for beginning instruction in reading and arithmetic on subsequent performance reported in the Plowden Report (9, Appendix 10).) In a sense it should be better than current grade of enrollment -- the latter is an outcome measure, not a background variable. Coleman asked sixth-graders, "What grade were you in last year?" but the responses appear not to have been used either by him or Mayeske as background variables. Questions were also asked to ascertain whether the student had attended kindergarten or nursery school. Mayeske (32, p. 377) computed averages of the criterion variable for various responses with respect to kindergarten attendance as shown in Table 14.

Table 14. Student averages of criterion variables, by kindergarten attendance -- Coleman data

Grade	Attended Kindergarten	Did Not Attend Kindergarten	Nonresponse	Percent of Variance Explained
3	51.3	48.2	45.2	3.2
6	51.7	48.3	38.1	6.8
12	51.7	48.1	41.7	3.8

Source: Mayeske (32, Appendix III)

A clear advantage for kindergarten is demonstrated for all three grades, although as usual, the inference to be drawn is not clear. Enrollment in kindergarten is, of course, likely to be significantly correlated with income, either of the family or of the community.

Differences associated with nursery-school attendance were so slight that they could easily be judged inconsequential.

The international studies have generally analyzed differences in age, but three of their population subgroups (I, II and III), like the NAEP age groupings, are students within a one-year age window. Thus, differences attributable to age tend to be trivial. Population IV, however, is the group in the last pre-college year (seniors in high school in the United States). Within this group there is some substantial difference in age, and Purves (41), for example, found a negative correlation of -0.22 between literature achievement score and age for the U.S., and he presumes that "the older student in school at Population IV would tend to be the less able student who had been held back for some reason or other."

The NAEP 17-year-old group is unlike Population IV, since age is held nearly constant. Those in age group 17 may be out of high school, either by graduation or dropout, may be seniors, juniors, sophomores, etc. Thus, their grade *attainment* is the variable, and that is an outcome which presumably could be used more effectively if the age of entry into first grade was known. Perhaps, as an intermediate variable, both current grade level and age of entry into first grade should be recorded.

It has been customary in most studies of educational achievement to record sex of student. In many studies, e.g., Mayeske's reanalysis of the Coleman data (32, p. 370), relatively small differences are noted. His criterion averages are as follows:

Grade	Male	Female	Nonresponse
3	49.5	50.7	45.2
6	49.4	50.7	43.3
12	50.6	49.5	43.9

The reversal in sign of the difference from third and sixth grade to twelfth grade is a phenomenon that has been frequently noted.

With respect to specific subject matter, however, some major sex differences have been noted. Comber and Keeves (7, p. 149) showed that at all ages males had an advantage in science scores, and that the advantage increased materially from age 10 to age 14 to age 18. The difference was greatest for physics and least for biology. NAEP found the same pattern for science exercises for age groups 9, 13 and 17 and found that for young adults the male advantage was even greater (Report 7, pp. 35-38).

Since sex is currently being ascertained by NAEP and since sex differences have been demonstrated, at least in exercises related to certain objectives, there seems to be no reason why recording of sex should be discontinued. There also seems to be no reason why the issue should be argued further in this study.

2.3.3 Family Structure

Family structure frequently has been used as a background variable to explain part of the variation in educational outcomes. Measures of family structure that have sometimes been used are absence of either or both parents, number of siblings, and birth order of the student. Sometimes if both parents are not present in the home the situation is described as a "broken home" and has been further classified by cause, such as divorce or separation, or death.

Coleman (5) asked the parental question as follows: *"Who is now acting as your father (mother)? If you are adopted, consider your adoptive father (mother) as your real father (mother)."* Response categories for father were as follows (with obvious modifications for mothers' responses):

(A) My real father, who is living at home
(B) My real father, who is not living at home
(C) My stepfather
(D) My foster father
(D) My grandfather
(F) Another relative (uncle, etc.)
(G) Another adult
(H) No one.

Mayeske (32) developed the criterion scales shown in Table 15 for the classes of response given above. Recall that these scale values are averages of the criterion for the categories given. It is clear that for third and sixth grades a real mother not living at home has a stronger negative association with the criterion than surrogates for the real mother living in the home. The pattern does not hold for twelfth graders or for the real father living away from home. Note, in general, the greater association for grade 6 than for grades 3 or 12.

Coleman also asked for:

Number of people in home
Number of brothers and sisters
Number of brothers and sisters older than respondent
Number of older brothers and sisters who left high school before finishing.

Scales (averages of the criterion-achievement composite) are shown in Table 16 for third, sixth, and twelfth grades.

Table 15. Scale values assigned by Mayeske to categories of response to questions concerning parents -- Coleman data

Category	Grade 3	Grade 6	Grade 12
Who acts as father?			
My real father, who is living at home	51.0	51.4	50.9
My real father, who is not living home	46.5	45.9	46.7
My stepfather	47.9	47.4	47.7
My foster father	44.4	43.5	45.7
My grandfather	44.7	43.0	42.9
Another relative (uncle, etc.)	46.9	43.6	45.0
Another adult	44.9	43.6	45.8
No one	46.4	45.5	46.9
Nonresponse	42.9	43.7	43.4
Percent of variance explained	4.4	6.6	3.3
Who acts as mother?			
My real mother, who is living at home	50.5	50.9	50.4
My real mother, who is not living at home	43.0	42.3	47.1
My stepmother	47.6	45.7	48.4
My foster mother	45.2	43.7	45.9
My grandmother	46.2	42.2	42.6
Another relative (aunt, etc.)	45.4	43.0	44.2
Another adult	47.5	40.4	44.0
No one	42.4	41.2	46.5
Nonresponse	42.6	42.8	44.3
Percent of variance explained	3.1	6.5	1.9

Source: Mayeske (32, Appendix III)

The low average scores of students in two-person homes probably reflects the effect of a broken home, i.e., only one parent present. For all three of the grades listed the performance of students with no brothers and sisters is below that of students with

Table 16. Scale values assigned by Mayeske to responses concerning family size -- Coleman data

	Grade		
Responses	3	6	12
Number of persons in the home			
2	44.5	44.8	49.3
3	52.0	51.0	51.0
4	52.8	52.4	51.3
5	52.0	52.4	51.0
6	50.5	51.0	50.0
7	48.5	49.4	49.3
8	47.6	47.5	47.1
9	45.9	46.2	46.0
10	45.1	43.9	44.8
11 or more	46.0	44.5	43.3
Nonresponse	43.9	42.9	43.8
Percent of variance explained	6.8	9.0	3.1
Number of brothers and sisters			
None	49.1	49.3	51.0
1	52.8	52.3	52.4
2	52.6	52.1	51.9
3	51.4	50.9	50.9
4	49.1	49.1	49.6
5	47.5	47.7	48.3
6	46.1	45.9	46.5
7	45.3	44.3	45.3
8	44.0	42.5	44.1
9 or more	43.4	41.5	42.4
Nonresponse	43.0	41.4	43.3
Percent of variance explained	10.0	8.5	8.1
Number of older brothers and sisters			
None			51.7
1			50.8
2			49.5
3			47.8
4			45.7
5			44.7
6			43.0

Table 16. Scale values assigned by Mayeske to responses concerning family size -- Coleman data -- Continued

Responses	Grade		
	3	6	12
7			43.2
8			42.5
9 or more			42.5
Nonresponse			43.4
Percent of variance explained			5.7
Number of older siblings who dropped out of school			
Have no older brothers or sisters			52.1
None			50.4
1			46.2
2			44.0
3			43.2
4			41.8
5			42.0
6			42.6
7			41.6
8 or more			42.9
Nonresponse			45.5
Percent of variance explained			6.8

Source: Mayeske (32, Appendix III)

one or two. However, performance drops off sharply with four or more brothers and sisters. It is known that family size is negatively correlated with income. The apparent drop in performance with increasing numbers of older brothers and sisters for twelfth-graders could be accounted for by the same influences that affect negatively the performance of students from large families. It is apparent, however, that there is strong association between performance and the number of older brothers and sisters who have dropped out before finishing high school, indicating the influence of intra-family correlations in ability, attitudes, and so on.

Bachman in Youth in Transition (1) considers number of siblings as well as whether the home is "broken" and whether the home is broken by death or by divorce, separation, or similar causes. His criterion is Quick Test scores rather than a measure of academic performance, but it is known that the Quick Test has moderately high correlation with various measures of academic performance. He has shown the following unadjusted means by various categories of intact or broken homes (Appendix F):

Home intact	109.27
Home broken by death	107.56
Home broken by divorce, etc.	104.06

The unadjusted correlation ratio between status of the home and Quick Test scores is 0.14, so that the variance accounted for is only about two percent. Thus, one is led to believe that any gain to be achieved by determining whether the home is broken will be small. It is also apparent that if status of the home is determined, homes broken by divorce, etc., form a separate class and that homes broken by death need not be separated from intact homes.

Bachman also shows the predictability of number of siblings, both unadjusted and after adjustment for socioeconomic level (a composite of father's occupational status, parents' education and family possessions). Average Quick Test scores are shown in Table 17. It is apparent that number of siblings contributes something to explanation of variance even after adjustment for socioeconomic level, although the measures used for adjustment are subject to error, and the adjustment may be only partially effective.

Another way of looking at the importance of number of siblings is through explained variance. Bachman shows the following percentages of explained variance:

Explained by socioeconomic level alone	19.8 percent
Explained by socioeconomic level and number of siblings	23.6 percent

Table 17. Mean scores on Quick Test, and Quick Test net of socioeconomic level, for each category of family size -- Bachman data

Family Size (Number of Siblings)	Mean Score on Quick Test	Quick Test Net of Socioeconomic Level
0	111.7	110.9
1	112.8	111.2
2	110.7	109.4
3	108.2	107.8
4	107.2	108.0
5	105.8	107.4
6	102.2	106.0
7 or more	97.8	102.5

Source: Bachman (1, Table E-4-9)

Explained by socioeconomic level plus number of siblings plus race	33.0 percent

Thus, number of siblings adds 3.8 percent to predictability after socioeconomic level and race adds another 9.4 percent. His analysis shows that number of siblings accounts for only 1.8 percent beyond that accounted for by socioeconomic level and race. One must keep in mind that Bachman's study covers a national sample of only tenth-grade boys, hence has limited capability for generalization.

2.3.4 Ability and Educational Achievement

While it is an accepted truism that ability affects educational achievement, the absolute demonstration of this phenomenon has yet to be accomplished. All studies attempting to show this relationship have used measures of ability that themselves have components of educational aehievement. For example, one of the individually-administered IQ tests taken as a standard, the Wechsler-Belleview, uses educationally-influenced measures in developing its index of ability, the Intelligence Quotient. As part of its verbal

IQ, for instance, it includes a subtest of information, a subtest of vocabulary, and a subtest of arithmetic reasoning, all measuring both innate aptitudes and learned experiences.

Further, the nonverbal IQ includes subtests of block design, figure manipulation and sequencing of pictures to tell the best story. These subtests are again influenced by both innate ability and educational or other experiences, as can be seen by the types of remedial materials used in programs such as Head Start. In these programs, emphasis is placed on "play" with form-manipulative materials to increase discrimination processes, play that is carried out in most middle and upper-class homes automatically.

Hanushek (20) in "The Value of Teachers in Teaching" presents a model for testing the effect of teachers and says:

> *Innate abilities* present probably the most difficult concept measure in the whole model. In fact, it is not well understood how innate abilities enter into the educational process, and there is considerable controversy over the role of innate ability in education. The only consensus that seems to exist in this area is that common IQ scores do an inadequate job of measuring innate abilities.

He then proceeds to analyze his data, using progress in achievement from an earlier grade as a way of "controlling" for differences in IQ. Unfortunately, NAEP does not have this alternative open to it, so we proceed with reviewing some attempts to use ability as a control, recognizing the shortcomings in the measurement of ability.[2]

There are mountains of literature on the predictability of educational outcomes (particularly grades and attainment) from measures of ability (IQ). Unfortunately, most of this literature is of little value to this investigation and we do not attempt to summarize it here. Instead, we review some results where ability (or IQ) has been considered as only one of several background factors used in the explanation of variations in outcomes.

---

[2] Errors in measurement are not unique to measurement of ability -- they are inherent in most background measurements.

Bachman (1) used a battery of eight ability tests on tenth grade boys as follows:

Matrices
Gates test of reading comprehension
Anagrams
Maze tracing
GATB - numerical
GATB - verbal
Hidden patterns
Job information test

He also used the Quick Test of Intelligence and an oral paragraph comprehension test and, after some analysis, decided to use the Quick Test scores as a measure of ability for subsequent analysis.

In earlier sections we have referred extensively to Bachman and have used the Quick Test as a measure of achievement. Here we use it in its more traditional sense as a measure of ability. This duality of roles is inconsistent but useful to the development of the measures of association of interest to the study.

A measure of academic achievement used by Bachman is grades and we consider first the association between ability (Quick Test scores) and grades with and without consideration of other background factors. The relevant data are summarized in Table 18. They show that eight background factors (without Quick Test) account for 11 percent of variation in grades and with Quick Test 18 percent of variation, hence an *added* 7 percent due to Quick Test. Quick Test alone would have accounted for 13 percent of variability. Note that Eta^2 in the Bachman studies is comparable with r^2 (the proportion of variance explained) in other studies cited.

Results were similar for predicting political knowledge. Thirteen percent of variation was accounted for uniquely by Quick Test as well as by all eight

Table 18. Regression analysis of predictability of grades from background factors -- Bachman data for tenth grade boys

Background Predictors	Eta	Eta^2
Socioeconomic level	0.26	0.065
Number of siblings	0.18	0.031
Broken home	0.10	0.011
Family relations	0.21	0.042
Religious preference	0.16	0.027
Family political preference	0.11	0.013
Community size	0.10	0.009
Race (5 categories)	0.10	0.009
Predicting from above 8 characteristics		0.114
Predicting from Quick Test alone	0.36	0.128
Predicting from Quick Test and other 8 characteristics		0.184
Added explanatory power of Quick Test		0.070

background factors. Another seven percent of explanatory power was provided by the Quick Test, making 20 percent of variance explained (after adjustment for degrees of freedom).

Hauser (21) analyzed data for nearly 17,000 white public school students in grades seven to twelve in Davidson County, Tennessee in 1957. He used two measures of academic achievement:

M = Stanford Mathematics Grade Equivalent
W = Stanford Reading Grade Equivalent

Using regressions within schools (refer to Section 1.3 for discussion) he found the percent of variation in each that could be accounted for by regression. Then he added IQ as an explanatory variable and redid the analysis. The analysis was accomplished by the method of path coefficients, but is more easily comparable with other results in this chapter if put in the following form:

	Achievement Test	
	M	W
Percent accounted for by:		
Father's occupation		
Father's education		
Number of siblings	4.1	7.1
Added percentage due to IQ	28.7	27.3
Total explained	32.8	34.4

Again, the contribution of intelligence is substantial. It should be noted, however, that restricting the study to white students in a single county restricts the total variance in student scores because of intra-class correlations. Therefore, the amount of variation explainable by other background factors will be reduced and the amount attributable to IQ will be proportionately greater. This is particularly true when the analysis is done within schools.

McDonald (35) used the Differential Aptitude Test in conjunction with SES, a measure of motivation (M-scale) in the prediction of grade point average. He found the following results for the "normal" population, i.e., neither underachievers nor overachievers:

	Sample (percent accounted for)	
	Validation	Cross Validation
Males		
SES and M-scale	0.34	0.35
Added by D.A.T.	0.34	0.34
Total	0.68	0.69
Females		
SES and M-scale	0.24	0.15
Added by D.A.T.	0.26	0.51
Total	0.50	0.66

Again, the contribution of ability is substantial.

Thus, in spite of the pessimistic note with which we began this section, we have shown that research has demonstrated some substantial predictability of performance from ability measures. Whether these measures do, in fact, measure ability is an unresolved issue, but may be a highly relevant issue as far as the interpretation of NAEP outcomes is concerned. That is, if a presumed ability measure actually measures outcome and NAEP adjusts its own outcome measures by the presumed measure of ability, it will actually be adjusting outcomes by other outcomes, making interpretation difficult.

Alan Wilson, in "Social Class and Equal Educational Opportunity" (Harvard Educational Review, Vol. 38, 1, Winter, 1968, pp. 77-84) (52) discusses the uses of ability as a controlled variable in studying achievement.

> In contrasting school environments it is clearly necessary to compare students who are similar in all relevant ways ... if differences are to be attributed to the contrasted schools which they attend ...
>
> Nevertheless, the reasons for not controlling for differences in performance on a concurrent ability or IQ test score when examining effects of environmental variations upon performance are compelling. Standard intelligence and ability tests are measures of specific knowledge and problem-solving skills which have been acquired by the testee at some time prior to the test situation. The validity of the IQ test score as a measure of learning "potential" depends upon the assumption of equal exposure to and practice with the kinds of knowledge and skills that the test calls upon. Since the tests were designed to predict performance in school, they call upon the kinds of knowledge and cognitive skills that are required in school. Thus the hypothesis under investigation must be assumed to be false in order for this control to be valid.
>
> One alternative design to circumvent this difficulty would be to undertake a longitudinal study comparing children with similar measured abilities early in their school careers who are subsequently exposed to contrasting school experiences. A study

using this strategy was conducted in one community after the publication of the Coleman Report and substantially confirmed the finding that inter-school differences do affect student achievement.

Our review of the literature is inadequate to put any kind of limits on the association of ability with achievement, independent of other factors, but Jencks (28) has speculated as shown below.

The available data suggest that:

1. If we could equalize everyone's genes, inequality in test scores would probably fall by 33 to 50 percent.
2. If we could equalize everyone's total environment, test score inequality would fall by 25 to 40 percent.
3. If we merely equalize everyone's economic status, test score inequality would fall by 6 percent or less.
4. Equalizing the amount of schooling people get might reduce cognitive inequality among adults by 5 to 15 percent, although this estimate is very rough.
5. Equalizing the quality of elementary schools would reduce cognitive inequality by 3 percent or less.
6. Equalizing the quality of high schools would reduce cognitive inequality by 1 percent or less.
7. Eliminating racial and socioeconomic segregation in the schools might reduce the test score gap between black and white children and between rich and poor children by 10 to 20 percent.
8. Additional school expenditures are unlikely to increase achievement, and redistributing resources will not reduce test score inequality.

It is obvious that there are substantial problems in separating performance and ability. These problems are accentuated when one considers population subgroups that have, for a variety of reasons, suffered some form of educational disadvantage. Further, the problems of administering any kind of uniformly interpretable ability tests within the NAEP testing environment are nontrivial. The recent Anchor and Equating test project may provide a partial solution, but there seem to be persuasive arguments against attempting to adjust individual exercise P values for differential abilities.

2.4 School Variables

2.4.1 Introduction

Many studies have investigated the relationship of school (or school district) variables to performance measures. In some studies school characteristics are treated as independent (or classification or regression) variables in an effort to understand the extent to which such variables, along with others, are associated with, or can "explain," *individual student performance*. In others, school (or district) averages or characteristics serve as the background or other independent variables to "explain" *average performance of the school or school district*. The motivation of studies of either type has sometimes been to examine the association between outcomes and segregation or integration and sometimes to identify various types of school or school district characteristics or practices that are consistently associated with high performance, in an effort to evaluate and improve educational practices and systems. Like other variables, some of the school (or district) variables are subject to school or educational system control and manipulation, and others are background variables that are not subject to control, at least on a short-run basis.

For some variables the school or the school district is the unit for which relevant information is available and consequently is the natural unit to be used. The school is a natural unit for such variables as pupil-teacher ratios, size of school, and many others. The district is a natural unit for factors related to finance: cost per pupil, average teacher salary, and so on, although such data are sometimes available or can be approximated for individual schools.

Another set of variables can be associated with either the school or the district, such as median income of family, percent of population in specified racial or ethnic groups, average educational attainment of the population, predominant occupations, and so on. If they can be shown to serve as useful background measures, characteristics of the community represent a potentially useful set of variables for NAEP. They often can be obtained from secondary sources such as the Census reports, the reports or records of the National Center for Educational Statistics, the Office of Civil Rights, and other sources. The National Center for Educational Statistics has created a school-district data base from the 1970 Census returns for small areas which can supply approximate Census data at the school district level. This data base should be especially useful in studies in which the school district is the unit of analysis.

Still another set of variables is associated with the student, but can be averaged over the students in a grade within a school or over the school district, or over the students in a given age or age group in a school or district, to create a school variable. An example of an outcome variable of this type is average achievement test score for all students in a particular grade level or other specified group in the school (or district). Examples of such background variables are average measures of pupil ability, average education or occupational score of pupils' parents, or average score for educational materials in the home. These variables may be related either to the specified grade or other subclasses within a school, or to the total school or school district. Note that the characteristic that distinguishes these variables from those in the previous paragraph is that here we are concerned with the average for the student body while in the previous paragraph we are approximating the attendance area. The attendance-area variables frequently can be obtained from secondary sources, while the student body variables must be obtained from the pupils themselves, or from their parents, or from school records.

Other school variables include the characteristics of a student's teacher in a subject or a class, the method of teaching in the particular class, and related variables often determined by the particular school class or classes that the student has taken or is taking. In this discussion we shall regard such variables as simply individual student characteristics, but because they are associated with the school we shall refer to them as student-linked school variables. As indicated earlier, our principal concern in this section is with school-wide variables, although some brief consideration will be given to student-linked school variables.

We can summarize our pragmatic classification of variables considered in this chapter as follows:

- School district variables -- expenditures per pupil, average (or median) teachers' salaries, average (or median) income, total pupils, average annual expenditures for plant and equipment, etc., for the school district

- School-linked (or school-wide) variables --

  Community-related -- average (or median) income, percent nonwhite, etc. for school attendance areas

  School characteristics -- school averages of student body characteristics, sex, age, and racial distribution of teachers, kind of school (academic/vocational, etc.) average characteristics of families of students, etc.

- Student-linked school variables -- Curriculum in which enrolled, characteristics of classroom or subject-matter teacher (of the individual student), etc.

One of the questions which arises in the adjustment of student outcome data is how specific to the individual student the background data must be in order to be useful for adjustment purposes. For example, in Section 2.2 it was shown that father's occupation and father's educational attainment were useful explanatory variables in the measurement of outcomes. One can ask the students in the upper grades and the teachers of students in the lower grades for this kind of information. However, for younger students, at least, one must assume there is some substantial error in student reporting of either occupation or education. Finally, there has been resistance in some jurisdictions to permitting students to report characteristics of their families on the grounds of invasion of privacy.

A key question is whether, if one did not have these parental characteristics for individual children, the average educational attainment of adults and the occupational distribution within the school attendance area would be meaningful codes to attach to the individual students' records for adjustment purposes. And, if such data are not available by attendance area, could school district or county data be used instead? Indeed, since occupation is used as a surrogate for income, might one reasonably use median income for the smallest geographic unit encompassing the attendance area for which such data are reported?

The answers depend upon the relative importance of community variables as contrasted to individual student background variables in the explanation of variance, and upon the proportion of the total variance in individual student outcome variables that is accounted for by the variability between school means. Not all of the answers will emerge from the analysis that follows, but in the analysis we will be aware of the need to answer such questions.

We now turn to a very brief consideration of student-linked school variables, and then to an examination of the use of school-linked and district variables in the analysis of individual student characteristics. Following that, we present the use of school-linked or district variables when the school or school district is the unit of analysis.

2.4.2 Student-linked School Variables

As indicated above, student-linked school variables relate to individual teacher, class, or other school characteristics associated with the individual student. Such variables presumably are subject to school or educational system control.

Few large studies have investigated the association between achievement and the within-school environment surrounding the individual student. Most large studies have considered the entire school as furnishing the environment within which the child learns. Since this literature review, by design, has given primary attention to large scale studies, we have undoubtedly missed some smaller scale, but important, studies of within-school relationships.

However, one of the student-linked school variables that has been extensively studied is the racial composition of the class (not the schools) in which the student is enrolled. For example, McPartland (36) in a study of 5,075 ninth grade Negro students in New England and the Middle Atlantic states showed that the racial composition of the class was an important factor in explaining individual student achievement, regardless of the racial composition of the school or the teaching system (if any) employed by the school.

Another student-linked variable that has received considerable attention is class size. Furmo and Collins (17) in a study of Baltimore public schools controlled on (1) parent's occupation, (2) intelligence test score, (3) type of curriculum and (4) race. They concluded that there was a significant advantage shown in favor of smaller classes. Studies using the entire school as the unit of analysis (e.g., the Coleman studies) have found less significance in class size.

A student-linked variable that we were surprised not to find in any of the large scale studies we reviewed is any characteristics of the classroom (or subject-matter) teacher that can be linked to the individual student. Mood (50) identified six categories of teacher variables as follows:

1. Dedication to the educability of all children
2. Ability to communicate
3. Ability to motivate
4. Ability to organize and manage a class
5. Ability to create learning experiences
6. Knowledge of a chosen field in which to teach

Variables of the kinds identified above related more to the process of education than to background. There are also nontrivial problems of measurement in trying to use them. Furthermore, a student's performance in a given year reflects his associations with teachers over a period of years, so that the characteristics of teachers in the entire school may, in fact, be a better measure of background (for the purposes of National Assessment) than characteristics of the specific teacher with whom the student has daily contact. This is not to say that it has equal analytic power.

2.4.3 School-linked (or School-Wide) Individual and Student Outcome Measures

It is the association of school-wide variables with individual student outcome measures that is of principal concern to NAEP.

Some analyses of special interest in connection with understanding and estimating the proportion of variance in individual student achievement or other outcome measures that can be associated with school-wide variables are given by Mayeske (31), by Marshall Smith (37), in Mosteller, Moynihan, et. al, by Purves (41) and by Husen (26).

Table 19, from Mayeske (31) shows, for Coleman study data, and for selected variables, the percent of the total individual student variance in achievement that is associated with the schools that students attend. This percentage represents the upper limit for the percent of achievement variance that can be associated with school-wide measures, either taken one at a time or through multiple regression. Thus, as indicated by Mayeske, no more than approximately 37 percent of the total achievement variance of the students in the ninth-grade level can be accounted for by correlating school-wide variables, such as school averages for student and teacher characteristics, with the individual achievement scores.

Table 19. Percent of total variance for different student variables that is associated with the schools students attend -- Coleman data

Variable Title	Grade Level 12	9	6	3
Socioeconomic Status	27.67	33.09	28.05	39.98
Family Structure and Stability	12.20	18.06	18.05	24.10
Racial-Ethnic Group Membership	68.97	68.58	55.54	59.62
Achievement	34.04	36.68	35.48	35.63
Attitude Toward Life	15.89	21.77	13.26	9.10
Expectations for Excellence	6.10	11.00	9.90	15.13
Educational Plans and Desires	10.12	11.26	12.38	9.92
Study Habits	11.31	18.40	15.04	19.41
Number of Schools	780	923	2,372	2,453
Number of Students	96,409	133,136	123,386	129,774

Source: Mayeske (31, Table 4.2)

Between-school variances are quoted by Marshall Smith, again with Coleman data, but using only the data for northern Blacks and Whites, and using the individual verbal achievement scores as the dependent variable. He gets the results shown in Table 20 for between-school variances as a portion of the total variance of individual achievement scores. These proportions are considerably lower than those from Mayeske in Table 19 as is to be expected in a restriction of the universe of analysis.

Table 20. Percent of total variance in individual verbal achievement scores that lies among schools*

	Level				
	12th Grade	9th Grade	6th Grade	3rd[+] Grade	1st Grade
Northern Blacks	10.92	12.67	13.89	19.47	10.63
Northern Whites	7.84	8.69	10.32	11.42	11.07

[*] From Table 3.22.1 of the EEOS Report.

[+] The 3rd-grade verbal test was unreliable (see Jencks, Chapter 4).

Source: Smith, in Mosteller and Moynihan (37, Chapter 6, Table 4)

Purves (41) reported for the U.S. sample the proportion of total student variance in literature scores accounted for by between-school variance as follows:

14-year-olds	25 percent
High school seniors	17 percent

The results are important here to illustrate the upper limit of what can be accounted for by school-wide variables in particular cases. Of course, the proportion of variance accounted for by between-school (or school-district) means is a function of the extent to which there is an intraclass correlation among students within a school on performance or outcome measures. The intraclass correlation is approximately equal to the proportion of variance that is accounted for by between-school variance.

The variance of individual student scores will be larger than that of school means, for any variable. Moreover, the variance of school means ordinarily will be larger than the variance of school district means, and these, again, will be larger than those of county means. This is not a mathematical statement, but it is commonly observed to be true and will be true in general when there is a positive intraclass correlation among students within a school with respect to whatever variables are under examination, and, similarly, when there is a positive intraclass correlation among schools within a school district, or among districts within a county. There are rational reasons to expect that these correlations will generally (although not universally) be positive. They will be positive whenever the units grouped into a higher-level unit (e.g., students within a school or schools within a district, or students within a district), tend to be more alike than any pair of units selected at random from the whole universe of such units. The results of many analyses suggest that such positive intraclass correlations do, in fact, describe what has been rather widely observed (17, Vol. I, p. 307).

We turn now to reporting some illustrative results of the proportion of total individual-student-score variance accounted for by school variables.

The Coleman study (5) focused directly on the issue of the association between school variables and outcomes within specific combinations of ethnicity and region. However, Coleman used verbal ability measures rather than a more direct measure of performance as a criterion or dependent variable, partly on the grounds that (1) between-school variances were maximized by this measure, and (2) between-school variances persist through the twelfth grade rather than decline as in the case of a performance measure. He points out that the verbal ability test is in fact a verbal achievement test and, in fact, uses "achievement" to denote it in the analysis. Even so, for our purposes, a more outcome-oriented measure of performance would seem to be closer to NAEP exercises than a measure of ability, and the decline in association through the upper grades (if this phenomenon in fact obtained for NAEP) should be reckoned with.

Coleman adjusted verbal ability scores of individual students for school background measures after first adjusting them for the following individual students' background factors:

Objective Measures

A. Urbanism
B. Parents' education
C. Structural integrity of the home
D. Smallness of family (number of brothers and sisters)
E. Items in home (TV, telephone, record player, refrigerator, automobile, vacuum cleaner)
F. Reading material in home (dictionary, encyclopedia, daily newspaper, magazines, books)

Subjective Measures

G. Parents' interest
H. Parents' educational desires

There are some differences in items included and sources of the data over the three grades studied, sixth, ninth, and twelfth. Most of the relevant analysis controls on the six characteristics A through F above. That is, subsequent analysis is done on the residual of the performance measure from the linear regression of performance on these six variables.

Table 21 shows the amount of the total student variance accounted for by the six individual background factors separately for specified subgroups of the population, and for each of three grades. There is evidence that associations with the student background factors decline from sixth to ninth to twelfth grade, possibly because of dropouts, thereby increasing homogeneity among those in the higher grades. It is also evident that the background factors are relatively unimportant in accounting for variation in outcome among northern Negroes and relatively important among southern Negroes and Indian-Americans.

The proportions of individual-student variance accounted for by school characteristics after adjustment for the six background variables given above are

Table 21. Percent of total student variance in verbal achievement accounted for at Grades 12, 9, and 6, by six individual student background factors -- Coleman data

	Grade		
	12	9	6
Puerto Ricans	3.64	3.89	23.71
Indian-Americans	18.89	13.92	18.40
Mexican-Americans	7.92	12.79	21.82
Negro, South	14.41	12.27	14.66
Negro, North	7.53	7.68	9.51
Oriental-Americans	11.81	12.75	34.77
White, South	14.75	18.40	18.14
White, North	14.28	16.49	14.10
Negroes, Total	13.48	12.15	14.01
Whites, Total	14.71	17.81	16.20

Source: Coleman (5, p. 300)

given in Table 22. The variability of the percentages over the population subgroups is considerable. Again, the association is relatively large for southern Negroes and Indian-Americans. However, note that the effect of school variables does not appear to decline from the third to twelfth grade, but rather to increase.

Coleman's assessment of the importance of school factors has been criticized (see Mosteller, Moynihan et. al. (37)) on the grounds that adjustment for student background factors first actually adjusts for part of the school factors, because of the correlation between student background and school factors. Marshall Smith (37, Chapter 6) gives an extensive analysis of the effects of three groups of school-wide variables, along with individual home background variables, in accounting for the variance of verbal achievement scores for sixth-grade Northern whites.

Table 22. Percent of variance in verbal achievement of individual students accounted for by school characteristics after control for individual student background characteristics -- Coleman Data

	Grade			
	12	9	6	3
Puerto Ricans	6.67	4.07	3.21	[illegible].27
Indian-Americans	11.48	2.59	5.64	[illegible].04
Mexican-Americans	6.59	2.82	1.47	3.50
Negro, South	8.64	7.52	4.90	0.80
Negro, North	3.14	1.45	0.77	2.96
Oriental-Americans	3.83	5.66	9.06	2.62
Whites, South	3.16	1.60	0.57	0.83
Whites, North	1.87	0.73	0.32	0.33
Negroes, Total	6.96	5.19	2.77	2.26
Whites, Total	2.53	1.15	0.47	0.33

School characteristics are:

Per pupil expenditure on staff
Volumes per student in library
Science lab facilities (9 and 12 only)
Extracurricular activities (9 and 12 only)
Presence of accelerated curriculum (9 and 12 only)
Comprehensiveness of curriculum (9 and 12 only)
Use of tracking (9 and 12 only)
Movement between tracks (9 and 12 only)
Size
Guidance counselors (9 and 12 only)
School location (city suburb, town, country)

Source: Coleman (5, p. 306)

Table 23 shows a list of variables used, and simple correlations with individual student verbal achievement, along with means and standard deviations for each variable. All of the variables except the home-background factor are school-wide variables (fuller definitions of the individual variables are given in (37) beginning on p. 321).

Table 23. Means, standard deviations, and standardized regression coefficients for selected independent variables in a number of regression equations with Verbal Achievement as the dependent variable - sixth grade northern Whites

Variable	Mean	S.D.	Zero-order r with Verbal
Home Background Factor			
3 Reading Material	0.293	0.510	0.23†
4 Items in Home	0.311	0.340	0.22†
5 Siblings	0.241	0.758	0.13†
6 Structural Integrity	0.236	0.580	0.11†
39 Parents' Education	0.277	0.877	0.28†
40 Urbanism of Background	0.067	1.242	-0.10†
Facilities and Curriculum Factor			
38 Per-Pupil Expenditure	492.614	174.[illegible]5	0.04
23 Volume Student	4.687	5.124	0.02
9 School Size	80.215	47.606	0.01
23 School Location	3.654	1.745	-0.00
37 Promotion Slow Learners	2.696	1.451	-0.03
36 Accelerated Curriculum	3.074	1.165	-0.03
Student Body Factor			
10 Proportion Own Encyclopedia	0.807	0.113	0.21†
26 Student Transfer	7.507	6.501	-0.00
31 Attendance	91.132	2.520	0.08†
55 T Perception of Student Quality	-0.135	0.423	-0.23†

Table 23. Means, standard deviations, and standardized regression coefficients for selected independent variables in a number of regression equations with Verbal Achievement as the dependent variable - sixth grade northern Whites -- Continued

Variable	Mean	S.D.	Zero-order r with Verbal
Teacher Characteristics			
11 T SES Level	3.729	0.634	0.07*
12 T Experience	11.736	4.561	-0.01
13 T Localism	0.055	0.796	0.07*
15 T Degree Received	3.022	0.317	0.05
16 T Preference for Middle Class	-0.240	0.712	-0.10†
17 T Verbal Score	23.341	2.063	0.05
18 T Prop. White	0.965	0.074	0.09†
48 Prop. School White	0.877	0.143	0.16†

* Significant at the 0.05 level

† Significant at the 0.01 level

Source: Smith (37, Chapter 6, Table 3)

Table 24 shows the percentages of total variance of individual verbal achievement scores that are accounted for by the three sets of school-wide groups of variables, showing the share of the total variance that is accounted for by each group of variables individually, and the joint contributions. The four objective background controls are: Home Items, Reading Items, Number of Siblings, and Structural Integrity of Family. The two subjective controls are: Parents' Educational Desires and Parents' Interest in Child's School Experience.

Table 24. Percentages of total variance of individual verbal achievement scores for three sets of schoolwide influences (Student Body, Facilities and Curriculum, and Teacher Characteristics) - sixth grade northern Whites

	With Background Controls		No Background Controls
	1 (4 obj., 2 subj.)	2 (6 obj.)	3
Student Body Uniquely	1.98	1.10	4.32
Facilities and Curriculum Uniquely	0.31	0.24	0.22
Teacher Characteristics Uniquely	0.55	0.46	0.53
Student Body and Facilities	0.00	0.01	0.16
Student Body and Teachers	0.96	0.41	2.01
Facilities and Teachers	0.07	0.08	0.00
Student Body and Facilities & Teachers	-0.09*	-0.09*	0.06
Total	3.78	2.22	7.35

* The negative signs indicate that the total variance explained by all three variables together is greater than the sum of the amounts of variance explained by the three variables separately and in pairs. This reflects the fact that some of the factors are negatively correlated with one another.

Source: Smith (37, chapter 6 Table 2)

In the *International Study of Achievement in Mathematics*, (26) as another illustration, a number of variables were used, as follows:

Parental Variables

Mother's education
Father's education
Father's occupation (status)
Father's occupation (scientific or nonscientific)
School standard deviation in Father's occupational status
Place of parents' residence

Teacher Variables (School-wide)

Student's opportunity of learning the test items
Description of mathematics teaching and school learning
Length of training
Sex of teacher
Recent in-service mathematical training
Degree of freedom given to the teacher

School Variables

Number of weekly hours of mathematics instruction
Number of hours in the school week
Number of hours of mathematics homework in the week
Number of hours of all homework in the week
Total roll of school
Percentage of men teachers on the school staff
Educational differentiation
Number of subjects taken in grade 12
Number of subjects taken in grade 8
Cost per student in teacher's salaries

Student Variables

Sex of student
Age of student
Student's level of mathematical instruction
Student's interest in mathematics

The "teacher variables" are based on the characteristics and perceptions as reported by the student's mathematics teachers. Some results from the study for these variables are summarized in Table 25. These data are for the national sample of U.S. students at the level indicated.

Table 25. Estimated simple correlations of individual teacher characteristics variables with mathematics test score

Variable	13-year-olds	High School Seniors with and without Mathematics Emphasis in High School: With	Without
Sex of teacher	0.01	-0.05	0.01
Length of training	0.08	0.15	0.03
Recent in-service mathematical training	0.10	0.18	0.02
Degree of freedom	0.03	0.03	0.00
Description of teaching	0.02	-0.02	-0.09
Student's opportunity of learning	0.19	0.29	0.04

Source: Husén (26) Vol. II, Tables 6.6, 6.8, and 6.9

It is seen that the simple correlations with mathematics test score are all relatively low, although they are somewhat higher for the seniors with mathematics emphasis in high school. This is particularly true for "length of training" and "recent in-service mathematical training." Also, the correlation of the teacher's report of student's advance opportunity of learning the mathematics content is relatively high for the 13-year-olds, and for the seniors with emphasis in mathematics.

Measures of the joint effect of these teacher variables are also given in the study, in the form of the proportions of the total variance (R^2) that is accounted for by this particular group of variables in the multiple correlations. These are given in Table 26.

Table 26. Percent of the variance accounted for by the group of teacher variables in Table 25 when analyzed in combination with a number of other variables

Student Population	Percent of Total Variance: With All Listed Variables (Plus Others) in Multiple Regression	Percent of Total Variance: Same but Excluding Student's Opportunity of Learning
13-year-olds	2.6	0.5
Seniors		
With Mathematics Emphasis	4.8	1.3
Without Mathematics Emphasis	3.8	2.8

Source: Husén (26) Vol. II, Tables 6.6, 6.8, and 6.9

It is important in interpreting these contributions to total variance to be aware that they come from a multiple regression analysis, and the contribution of these particular variables will be a function of what other variables are also included in the regression analysis. The other variables include:

a. Five variables representing parents' characteristics and including mother's and father's education, father's occupation (and whether it is a scientific occupation or not), and place of residence of parents.

b. Six student variables including sex and age, student's level of mathematical instruction, and student's interest in mathematics.

c. Ten school variables including averages for weekly hours of mathematics instruction, hours in the school week, school size, percentage of teachers male, number of subjects in grade 12, number of subjects in grade 8, and cost of teachers' salaries per student.

It should be especially emphasized that Table 26 does not represent the added contribution of these variables to variance accounted for, in a stepwise regression approach, but rather their share[3] of the total variance accounted for in the particular regression equation. Multiple regression and multiple correlation measures of variance contributions for subsets of variables are difficult to interpret because results for any particular variable or subset of variables depend on the other variables involved, and in some instances, on the order in which the variables are introduced. Usually the square of the simple correlation coefficient between a specified variable and the dependent variable shows the upper limit of what the specified variable can account for in proportion of variance, no matter what other variables are introduced, although in certain special cases this may not hold.

The simple correlations of the specified schoolwide variables with mathematics scores are given in Table 27. It is seen that only a few of the correlations are of any consequence in terms of their potential for accounting for variance by themselves. The fraction of total individual-student mathematics score variance accounted for in a multiple regression invoking all variables listed earlier is given in Table 28. It is seen that school-wide variables are more effective for seniors than for 13-year-olds, and apparently more effective for the mathematics seniors than for the other seniors. Note, however, that the school-wide variables are tailored for their association with mathematics scores. The proportions of variance accounted for in the multiple regression by all the school-wide variables combined is not large.

---

[3] Obtained by summing br for the specified variables from the regression equation, where r is the simple correlation coefficient of the indicated variable with the dependent variable, and b is the corresponding standardized regression coefficient. See (26), p. 261.

Table 27. Simple correlations of school-wide variables with individual-student mathematics scores, for the U.S. sample

	Simple Correlations For		
School-wide Variables	13-year-olds	Mathematics Students	Other Students
Time for all schooling	-0.04	-0.16	-0.14
Time for all homework	0.14	0.22	0.17
Time for instruction in mathematics	0.02	-0.22	-0.08
Time for mathematics homework	-0.01	0.09	0.05
Total role of school	0.06	0.32	0.27
Percentage of male teachers	0.01	0.04	0.07
Number of subjects in grade 8	-0.03	0.01	0.01
Number of subjects in grade 12	0.01	0.10	-0.14
Cost per student (teachers' salaries)	0.10	0.07	0.08
Educational differential	-0.01	-0.18	-0.16

Source: Husén (26) Vol. II, Tables 6.14, 6.16, 6.17

Mayeske in A Study of the Achievement of Our Nation's Students (33), again using Coleman data, has shown that about 48 percent[4] of the variation in

---

[4] Figures are read from Mayeske's graphs and are subject to some error.

Table 28. Fraction of total individual student mathematics score variance accounted for by specified groups of school-wide variables, in a multiple regression with additional variables, for the U.S. sample

Groups of School-wide Variables	13-year-olds	Seniors	
		Mathematics Students	Other Students
First group of four variables listed in Table 27	2.1	3.6	1.3
Other school-wide variables	0.4	6.2	5.3
All school-wide variables	2.5	9.8	6.6

Source: Husén (26) Vol. II, Tables 6.14, 6.16, 6.17

individual student achievement can be accounted for by home background, race (ethnic group), and a motivational factor which he identifies as Family Process. Home background includes socioeconomic status and family structure and stability. Family Process includes students' reports of parents' expectations for excellence, attitude toward life, educational plans and desires, and study habits. About 37 percentage points of the 48 percent accounted for by home background, race and family process are accounted for by home background and race alone.

If we take the approximately 37 percent of variance accounted for by home background and race/ethnicity as a base, the key question as far as this study is concerned is how much additional explanation in variation of *individual student achievement* can be obtained by adding school factors. Unfortunately, Mayeske's analysis does not provide exactly the answer we need. First, he added family process to home background, so that we work from a base of 48 percent rather than 37 percent. Second, he included in school variables average student body (1) expectations for excellence, (2) attitude toward life, (3) educational plans and desires, (4) study habits, and (5) achievement. The first four of these variables are included in family background, as individual student characteristics, and the fifth is the average of the criterion for the school. Thus, the amount of additional variance explainable by school factors (as contrasted to average student-body characteristics) is difficult to ascertain.

However, he shows that the total variation explained was 54 percent, so that the additional amount explained by school factors must be no greater than 54-48, or 6 percent. Note again that this does not argue that school factors are unimportant since, as demonstrated earlier, they are highly interlocked with student characteristics.

Mayeske has been able to explain more of the total student variation than most other major studies. He sorted through a large number of possible background variables and used criterion scaling. Both techniques tend to produce levels of association that cannot be achieved on an independent set of data. However, the smallest racial/ethnic group studied by him was Orientals and even for that group he had 1,675 students in the ninth grade, approximately equally divided by sex. Indians were represented by 2,877, Mexican-Americans by 5,836, Puerto Ricans by 3,702, and Negro by 37,265 ninth graders. He covered first, third, sixth, ninth and twelfth grades, so for some analyses the samples were about five times as large as the above figures. Thus, while his figures might not be duplicated by an independent set of data, they must be taken seriously as representations of the predictability that can be achieved with a large number of background variables.

One of the principal criticisms of the study is that the data are old, having been collected prior to the massive efforts to balance racial groups within the public schools. One must presume that this effort would have some effect upon the explanatory power of, say, racial distribution as a background variable. It must also have had some effect upon the homogeneity of school attendance areas and even the ability to define such areas.

Before leaving the Coleman study and its derivatives, the simple correlations found by him between "verbal scale" and selected school factors may be worth noting. They are summarized for three racial-ethnic groups and for grades 6, 9, and 12 in Table 29. The variables are described more completely in the Coleman report (5) and its supplemental appendix (6). Table 29 demonstrates the variety of schoolwide variables considered and the relatively low predictability of any one of them above. Complete correlation matrices are given in (6); it is not feasible to reproduce them here.

Table 29. Simple correlations between individual verbal score and selected school variables reported by Coleman

Grade and Variable	Mexican-American	Negro	White Majority
Grade 12			
Teachers' SES level	0.12	0.19	0.05
Teachers' experience	-0.02	0.13	0.00
Teachers' location	0.06	0.02	0.06
Quality of college teachers attended	0.20	0.23	0.05
Degree received by teacher	0.12	0.25	0.12
Teachers' preference for middle class student	-0.23	-0.15	-0.15
Teachers' verbal ability score	0.25	0.31	0.04
Teachers' race	0.34	0.23	0.01
Teachers' salary	0.12	0.28	0.13
Teachers' number of absences	-0.10	0.07	0.04
Pupils per teacher	-0.12	-0.11	-0.05
Science lab facilities	0.14	0.19	0.07
Volumes of books per student	0.00	0.06	-0.02
Extracurricular activities	0.17	0.20	0.12
Comprehensiveness of curriculum	0.09	0.09	-0.03
Student transfers in and out	-0.01	0.06	0.01
Movement between tracks	0.10	-0.03	0.03
Days in session	0.01	0.15	0.05
Teacher turnover	0.01	0.00	-0.03
Guidance counselors	0.13	0.28	0.15
Attendance	0.12	0.05	0.03
College attendance of last year's graduates	0.22	0.22	0.20

Table 29. Simple correlations between individual verbal score and selected school variables reported by Coleman -- Continued

Grade and Variable	Mexican-American	Negro	White Majority
School location	0.13	0.25	0.10
Length of academic day	-0.01	-0.04	0.02
Tracking	-0.03	-0.02	-0.03
Accelerated curriculum	-0.12	-0.11	-0.12
Promotion of slow learners	0.12	-0.02	-0.06
Per pupil instruction expenditures	0.09	0.21	0.10
Grade 9			
Teachers' SES level	0.16	0.18	0.05
Teachers' experience	-0.01	0.02	-0.06
Teachers' location	0.07	0.05	0.07
Quality of college teachers attended	0.21	0.21	0.02
Degree received by teachers	0.06	0.20	0.06
Teachers' preference for middle class student	-0.21	-0.13	-0.07
Teachers' verbal ability score	0.21	0.25	0.02
Teachers' race	0.36	0.22	0.06
Teachers' salary	0.09	0.24	0.11
Teachers' number of absences	-0.06	0.07	0.04
Pupils per teacher	-0.06	-0.06	-0.06
Science lab facilities	-0.08	-0.05	-0.04
Volumes of books per student	0.03	0.03	0.04
Extracurricular activities	0.07	0.05	0.07
Comprehensiveness of curriculum	0.03	0.05	-0.05
Student transfers in and out	0.05	0.04	0.01
Movement between tracks	-0.00	-0.11	-0.01
Days in session	0.02	0.13	0.04
Teacher turnover	-0.01	0.05	0.00
Guidance counselors	0.07	0.20	0.11
Attendance	0.13	0.07	0.05
College attendance of last year's graduates	0.16	0.20	0.13
School location	0.14	0.19	0.10
Length of academic day	0.05	-0.02	0.06
Tracking	-0.02	-0.06	-0.02
Accelerated curriculum	-0.10	-0.09	-0.06

Table 29. Simple correlations between individual verbal score and selected school variables reported by Coleman -- Continued

Grade and Variable	Mexican-American	Negro	White Majority
Promotion of slow learners	0.09	0.08	0.07
Per pupil instruction expenditure	0.07	0.19	0.11
Grade 6			
Teachers' SES level	0.24	0.15	0.08
Teachers' experience	-0.07	-0.03	-0.06
Teachers' location	0.22	0.01	0.08
Quality of college teachers attended	0.22	0.19	-0.00
Degree received by teachers	-0.03	0.04	0.05
Teachers' preference for middle class student	-0.15	-0.12	-0.08
Teachers' verbal ability score	0.31	0.24	0.07
Teachers' race	0.41	0.21	0.10
Teachers' salary	0.16	0.23	0.12
Teachers' number of absences	0.01	0.08	0.04
Pupils per teacher	-0.07	-0.04	0.00
Science lab facilities	-0.12	-0.08	-0.03
Volumes of books per student	0.11	0.03	0.02
Extracurricular activities	-0.06	-0.08	-0.03
Comprehensiveness of curriculum	0.03	0.22	*
Student transfers in and out	0.10	0.11	0.01
Movement between tracks	-0.06	-0.07	-0.01
Days in session	-0.08	0.07	0.02
Teacher turnover	0.05	0.06	-0.00
Guidance counselors	-0.06	0.01	-0.03
Attendance	0.28	0.16	0.07
College attendance of last year's graduates	0.04	-0.01	-0.01
School location	0.05	0.18	0.05
Length of academic day	-0.03	-0.09	-0.06
Tracking	-0.03	-0.03	-0.00
Accelerated curriculum	-0.09	-0.06	-0.04
Promotion of slow learners	-0.02	-0.02	-0.02
Per pupil instruction expenditure	0.10	0.22	0.10

* Obvious error in printout.

Source: Coleman (6, Supplemental Appendix)

2.4.4 School-wide Variables and School-Average Outcome Measures

In the above analysis we have focused on the student as the unit of analysis and have presumed that the ability to explain variation in individual student scores comes close to NAEP's objectives for the use of background factors. While NAEP data are reported by subgroups of the population, these subgroups are comprised of individual students. If their exercise scores can be adjusted for background individually it follows that adjusted subgroup P values can be computed.

An important part of the literature, however, has dealt with the problems of evaluation of performance of schools, or school districts, through regression or production function studies, or through simpler analytic comparative or correlation procedures. These studies do not appear to serve NAEP's purposes directly, but they do have relevance in that the school-wide variables ordinarily are of the same type as those in which there is interest in analyzing individual student performance, and these studies may shed additional light on the evaluation of these variables. Under many circumstances (but not all) it may be reasonable to assume that background or other variables that are more or less important in school-wide studies will have been ordered in about the same way, in terms of their relative importance in explaining variance of individual student outcome measures, even though the total amount of variance explained in individual outcome measures is typically considerably smaller.

We saw earlier, in Section 2.4.3 and Tables 19 and 20, some illustrations of the proportion of variance of individual student achievement scores that is accounted for by the variance between school means. It seems reasonable to assume that the information from regression or correlation studies at the school or school-district level does provide some rough boundaries or limits for interpretation of the contribution of school-wide variables to the variance of individual student outcome measures. (See Section 1.3 for a discussion of the relationship between residual variance from school-mean-regression and individual-student-regression.)

In any case, it has seemed desirable to comment here on some of the principal school and district studies, keeping in mind that here when we speak of explaining variation we are talking about the variation between schools (or districts) rather than the variation among students.

Mayeske (31, p. 9) created an individual student-achievement composite from tests of nonverbal and verbal ability (all grades), plus reading comprehension and mathematics achievement (grades 12, 9, 6, and 3), plus general information (grades 12 and 9). He constructed a Student Body Index from socio-economic status, family structure and racial-ethnic composition and a School Index which will be described in more detail later. His analysis of the associations among the individual student achievement composite for the school, student body, and school factors is shown in Table 30.

Table 30. Summary of Mayeske's analysis of the associations between student body and school characteristics and average school achievement -- Coleman data

Grade	R^2	Unique Components		
		Student Body	School	Joint
1	0.52	0.05	0.09	0.38
3	0.56	0.07	0.04	0.45
6	0.83	0.10	0.04	0.69
9	0.87	0.11	0.05	0.71
12	0.86	0.08	0.04	0.74

Source: Mayeske (31, p. 43)

The data presented in Table 30 are interesting because of the relatively high proportion of school variation in achievement that can be explained by school and student-body characteristics -- a proportion which increases dramatically from first to twelfth grade -- as well as for the evident commonality of the effect of student body and school. Student body factors predicted 82 percent of the between-school variation of twelfth-graders and school factors predicted 79 percent, but their unique contributions are only 5 and 9 percent, respectively. Thus, it is evident that adjusting for student-body characteristics does in fact adjust for most of the school effect as well. Again, one must keep in mind that these are percentages of the between school variation and not percentages of the total variation among students.

Mayeske, by multiple regression analysis, identified 31 school variables that comprise the school factor whose association with achievement is discussed above. The 31 variables are listed in Table 31. That table also gives partial correlations between some of the 31 variables and school achievement with school size, home background and race held constant (which accounts for the zero values shown for enrollment in Table 31). Apparently, pupil-teacher ratios, specialized staff and services, and teacher's vocabulary score have consistent association with achievement across all grades. It is not unlikely that all three are correlated with the financial support provided by the community.

Table 31. Partial correlations of Mayeske's school variables with achievement after control on school size, home background and race -- Coleman data

	Grade			
Variable	3	6	9	12
Facilities				
1. Plant and physical facilities				
2. Instructional facilities				
3. Age of building				
4. Pupils per room				

Table 31. Partial correlations of Mayeske's school variables with achievement after control on school size, home background and race -- Coleman data -- Continued

Variable	Grade 3	Grade 6	Grade 9	Grade 12
Pupil Programs and Policies				
1. Tracking				
2. Testing				
3. Transfers				
4. Remedial programs				
5. Free milk and lunch programs				
6. Accreditation				
7. Age of texts				
8. Availability of texts	0.04	0.07	0.08	0.08
9. Pupil-teacher ratio	-0.08	-0.14	-0.30	-0.06
10. Enrollment	0.00	0.00	0.00	0.00
School Personnel and Personnel Expenditures				
1. Principal's experience				
2. Principal's training				
3. Principal's college attended				
4. Principal's sex				
5. Principal's estimate of school's reputation	0.08	0.01	0.03	0.05

Table 31. Partial correlations of Mayeske's school variables with achievement after control on school size, home background and race -- Coleman data -- Continued

Variable	Grade 3	Grade 6	Grade 9	Grade 12
6. Specialized staff and services	0.10	0.10	0.23	0.18
7. Teacher's experience				
8. Teacher's training				
9. Teacher's socioeconomic background	-0.04	-0.01	0.04	-0.05
10. Teacher's localism				
11. Teacher's college attended	0.01	0.10	0.05	0.03
12. Teaching conditions	0.14	0.11	0.01	0.21
13. Teaching-related activities	0.01	0.01	0.01	0.05
14. Preference for student ability level	0.04	0.01	-0.03	-0.02
15. Teacher's sex				
16. Teacher's racial-ethnic group membership	0.01	0.18	0.17	-0.02
17. Teacher's vocabulary score	0.13	0.25	0.21	0.07

Source: Mayeske (32, Appendix XI).

Further analysis by Mayeske revealed the unique portions of explained variance associated with each of the three groupings of school variables listed in

Table 31. This analysis is equivalent to partitioning the unique contribution of the "school" and "joint" columns of Table 30. The results are shown in Table 32.

Table 32. Summary of Mayeske's analysis of the unique association between achievement and student body characteristics and groupings of school variables -- Coleman data

Unique Components	Grade 3	 6	 9	 12
First Order				
Student body alone	0.07	0.10	0.11	0.08
Facilities alone	*	*	*	*
Personnel and personnel expenditures alone	0.03	0.03	0.02	0.02
Pupil programs and policies alone	*	0.01	0.01	0.02
Second Order				
Student body and personnel	0.35	0.55	0.49	0.45
Others	0.01	0.01	0.02	0.01
Third Order				
Student body, personnel and programs	0.08	0.12	0.12	0.07
Student body, personnel and facilities	*	0.02	0.03	0.17
Others (including fourth order)	0.02	*	0.07	0.05
R^2	0.56	0.83	0.87	0.86

* Less than 0.005.

Source: Mayeske (31, p. 51)

It may be seen that the interaction of average student-body characteristics and school personnel dominates the explanation of between-school variance. Facilities do not enter with any significance except for twelfth grade, and then only in interaction with student body and personnel. Pupil programs and policies alone contribute a very small amount and join with student body and school personnel factors in contributing a nontrivial amount to the explanation of variance.

The California State Testing Program of 1970-71 (4) provides a good example of predictability of background factors at the district level. Twenty-two background variables were examined within elementary, unified, and high school districts with respect to their ability to predict reading scores for grades 1, 2, and 3 and reading, language, spelling and mathematics scores for grades 6 and 12. A stepwise regression program was used and the order of entry of the predictor variables is presented in Tables 33 through 35 for the three kinds of districts studied.

It is interesting to note that the first and second variables to enter the prediction equation in almost every case are variables which reflect poverty and minority enrollments. At the bottom of the tables are shown the proportions of variance in average district scores accounted for by all of the variables in the table, by the first one to enter the equation, and by the first two. It is apparent from these data that a major portion of the explained variance is accounted for by the index of family poverty and by minority enrollments.

One unusual feature of the California study is that the percent of variance explained by linear regression was cross-validated on random halves of the sample. The results of the cross-validation for unified school districts is shown in Table 36. These cross-validation results are not as favorable as for elementary districts but more favorable than for high school districts. They provide a means for judging repeatability of the results on an independent study. The sample sizes here, in terms of numbers of districts, are relatively small, and consequently overestimates of explained variance from a single sample will be considerably greater than would occur for much larger samples.

Extensive tables of regression and correlation results are presented in convenient reference form in the study.

Table 33. Ranks of predictor variables in terms of order of entry in stepwise regression program, California State Testing Program, 1970-71

Unified Districts

Predictor Variable	Order of Entry by Grade Level and Output Variable						
	Grades						
	1	2	3	Grade 6			
	Reading			Reading	Language	Spelling	Mathematics
20. Assessed Valuation/ADA	11	9	11	5	5	5	8
21. Percent Total Minority Pupils	6	6	5	4	11	10	5
22. Percent American Indian	10	11	9	8	6	9	7
23. Percent Negro	7	10	10	9	8	7	11
24. Percent Oriental	3	4	3	3	2	2	3
25. Percent Spanish Surnamed	2	3	2	2	3	3	2
26. Index of Family Poverty	1	1	1	1	1	1	1
29. Pupil Mobility Grades 1-8	4	5	6	7	4	4	4
30. Pupil Mobility Grades 9-12	-	-	-	-	-	-	-
31. Rate of Staff Turnover	8	7	7	11	9	11	10
32. Expenditures/ADA	5	2	4	10	10	8	6
33. Regular ADA Grades 1-12	9	8	8	6	7	6	9
R^2, First Variable Entered	0.20	0.32	0.38	0.36	0.36	0.30	0.34
R^2, First Two Variables	0.20	0.37	0.43	0.44	0.42	0.36	0.40
R^2, All Variables	0.33	0.46	0.56	0.55	0.59	0.49	0.56

Source: (4)

Table 33. Ranks of predictor variables in terms of order of entry in stepwise regression program, California State Testing Program, 1970-71 -- Continued

Unified Districts

Predictor Variable	Order of Entry by Grade Level and Output Variable			
	Grade 12			
	Reading	Language	Spelling	Mathematics
20. Assessed Valuation/ADA	4	6	4	4
21. Percent Total Minority Pupils	2	8	11	2
22. Percent American Indian	10	2	2	5
23. Percent Negro	11	9	9	7
24. Percent Oriental	3	4	5	3
25. Percent Spanish Surnamed	9	3	7	11
26. Index of Family Poverty	1	1	1	1
29. Pupil Mobility Grades 1-8	-	-	-	-
30. Pupil Mobility Grades 9-12	6	5	8	6
31. Rate of Staff Turnover	8	11	10	10
32. Expenditures/ADA	5	10	6	8
33. Regular ADA Grades 1-12	7	7	3	9
R^2, First Variable Entered	0.32	0.31	0.20	0.35
R^2, First Two Variables	0.39	0.35	0.27	0.38
R^2, All Variables	0.45	0.45	0.37	0.50

Table 34. Ranks of predictor variables in terms of order of entry in stepwise regression program, California State Testing Program, 1970-71

	Elementary Districts						
	Order of Entry by Grade Level and Output Variable						
	Grades						
	1	2	3	Grade 6			
Predictor Variable	Reading			Reading	Language	Spelling	Mathematics
20. Assessed Valuation/ADA	8	10	8	8	7	7	9
21. Percent Total Minority Pupils	1	1	1	1	3	8	4
22. Percent American Indian	10	11	11	11	10	11	7
23. Percent Negro	11	7	9	10	9	9	10
24. Percent Oriental	5	2	2	2	2	2	3
25. Percent Spanish Surnamed	7	9	6	7	8	3	-
26. Index of Family Poverty	3	3	5	3	1	1	1
29. Pupil Mobility Grades 1-8	6	5	7	5	6	4	2
30. Pupil Mobility Grades 9-12	-	-	-	-	-	-	-
31. Rate of Staff Turnover	9	8	10	9	11	10	8
32. Expenditures for ADA	4	6	3	4	4	5	6
33. Regular ADA Grades 1-12	2	4	4	6	5	6	5
R^2, First Variable Entered	0.07	0.11	0.14	0.17	0.12	0.12	0.14
R^2, First Two Variables	0.12	0.16	0.23	0.26	0.18	0.19	0.22
R^2, All Variables	0.18	0.26	0.33	0.36	0.29	0.26	0.26

Source: (4)

Table 35. Ranks of predictor variables in terms of order of entry in stepwise regression program, California State Testing Program, 1970-71

High School Districts

Predictor Variable	Order of Entry by Grade Level and Output Variable, Grade 12			
	Reading	Language	Spelling	Mathematics
20. Assessed Valuation/ADA	3	10	8	10
21. Percent Total Minority Pupils	11	8	4	2
22. Percent American Indian	8	5	3	7
23. Percent Negro	4	7	2	4
24. Percent Oriental	7	4	5	3
25. Percent Spanish Surnamed	1	3	10	9
26. Index of Family Poverty	2	1	1	1
29. Pupil Mobility Grades 1-8	-	-	-	-
30. Pupil Mobility Grades 9-12	5	9	9	5
31. Rate of Staff Turnover	9	6	-	8
32. Expenditures for ADA	10	11	7	6
33. Regular ADA Grades 1-12	6	2	6	11
R^2, First Variable Entered	0.48	0.38	0.25	0.44
R^2, First Two Variables	0.58	0.43	0.31	0.53
R^2, All Variables	0.60	0.51	0.37	0.58

Source: (4)

Table 36. Results of cross-validation study, unified districts

	Proportion of Variance Accounted for (R^2)			
Dependent Variable	Total Sample	(A) Average R^2 for Odd-Even Subsamples	(B) Average R^2 from Cross-Validation	Difference (A-B)
1 Reading	0.3262	0.3623	0.2095	0.1528
2 Reading	0.4615	0.4797	0.4199	0.0598
3 Reading	0.5609	0.6000	0.4754	0.1246
6 Reading	0.5559	0.6051	0.4397	0.1654
6 Language	0.5895	0.6015	0.5254	0.0761
6 Spelling	0.4947	0.5358	0.3363	0.1995
6 Mathematics	0.5636	0.5880	0.4467	0.1413
12 Reading	0.4462	0.4832	0.3153	0.1679
12 Language	0.4505	0.4616	0.3512	0.1104
12 Spelling	0.3661	0.3837	0.3101	0.0736
12 Mathematics	0.4965	0.4932	0.4481	0.0451

Source: (4)

Garms (18) studied the relationship between socioeconomic characteristics and pupil achievement in elementary schools in New York State. The school was the unit of analysis and he examined 303 SES variables. Schools were stratified into five strata: New York City, other urban, New York SMSA, upstate SMSA and non-SMSA. About 20 third-grade pupils were sampled in each school. Two basic outcome measures were obtained on the sample of students in each school:

- Average score on the Reading Test of the New York State Pupil Evaluation Program

- Average score on the Arithmetic Test of the New York State Pupil Evaluation Program

The principal outcome variables used in the analysis were percent scoring below the fourth stanine in each of the above tests and in the combined test. The 15 background variables most highly correlated with the above criterion variables are shown in Table 37.

Table 37. Simple correlations of Arithmetic and Reading test criteria with 15 most highly correlated background variables -- New York State data

	Simple Correlations with Percent Below 4th Stanine		
Background Variables	Reading and Arithmetic	Reading	Arithmetic
1. Average years of education of mothers	-0.677	-0.630	-0.658
2. Percent of children eligible for free-lunch program	0.676	0.610	0.677
3. Average years of education of fathers	-0.671	-0.627	-0.649
4. Percent of fathers with 10 or fewer years of education	0.660	0.616	0.639
5. Percent of children from broken homes	0.660	0.613	0.643
6. Percent of heads of households whose income is less than $5,000 per year	0.632	0.580	0.621
7. Percent whose occupation is other than professional, skilled, service worker, or farm worker	0.613	0.570	0.596

Table 37. Simple correlations of Arithmetic and Reading test criteria with 15 most highly correlated background variables -- New York State data -- Continued

	Simple Correlations with Percent Below 4th Stanine		
Background Variables	Reading and Arithmetic	Reading	Arithmetic
8. Percent living in home owned by parents	-0.613	-0.575	-0.591
9. Percent living in an apartment	0.590	0.538	0.585
10. Percent of fathers with 7 or fewer years of education	0.590	0.551	0.570
11. Percent of over-crowded housing	0.577	0.534	0.563
12. Percent Puerto Rican	0.551	0.524	0.523
13. Percent of low-rent apartments	0.549	0.508	0.537
14. Percent Negro	0.543	0.515	0.518
15. Percent whose father's occupation is professional	-0.529	-0.506	-0.499

Source: Garms (18)

The data are remarkable for their consistency across the two criterion measures and for their similarity among the 15 background factors. Also, it may be important to note that the measures (or a close approximation for them) may be obtained for the school attendance area from Census data.

Garms found that in using 30 SES variables he could account for 65 percent of the variation among schools in the percent of students below the fourth stanine in reading and arithmetic. With only three of the SES variables (percent in broken homes, percent in overcrowded housing and years of occupation of mother), including interactions among them, he was able to account for 62 percent of the variation. Thus the other variables accounted for only an additional three percent. The study verified results of an earlier New York study by Garms et. al.

The Fleischmann Report (15), which relied on some of the analytical work of Garms, found essentially the same results for third, sixth, and ninth grades. Some key results are shown in Table 38. They relate the distribution of the Regents Scholarship and College Qualification Test (RSCQT) to five SES categories based upon Garms' variables.

Table 38. Distribution of RSCQT scores by SES classes -- Fleischmann data

RSCQT Scores	SES I	SES II	SES III	SES IV	SES V
Upper quarter	50%	37%	31%	21%	13%
Upper half	26%	30%	27%	23%	18%
Lower half	17%	22%	25%	26%	27%
Lower quarter	7%	12%	17%	30%	42%

Source: Fleischmann Report (15)

A number of state studies show results that are not inconsistent with the above. Among the studies reviewed in the preparation of this report that show regression and multiple correlation measures are the following:

A Regression or Multiple-Correlation Study

H. J. Kiesling, "The Relationship of School Inputs to Public School Performance in New York State," P-4211, The Rand Corporation, October 1969 (29)

Other Analytical-Type Studies

Donovan, David and others, "Distribution of Educational Performance and Related Factors in Michigan." The Sixth Report of the 1970-71 Michigan Educational Assessment Program, Michigan State Department of Education, Research, Evaluation and Assessment Services, Lansing, Michigan, June 1972 (10)

"1971 Texas Achievement Appraisal Study," Texas Education Agency, Austin, Texas, May 1972 (46)

"Sixth Grade Reading: A Needs Assessment Report," Texas Education Agency, Austin Texas, 1972 (47)

Firman, William D., University of the State of New York, The State Education Department, "The Quality Measurement Project in New York State," presented at American Association for the Advancement of Science, Berkeley, California, December 29, 1965 (12)

Also, "Performance Indicators in Education-PIE," The University of the State of New York, The State Education Department, (51) includes a general description of issues in such studies, and a survey of 80 studies.

We believe that the studies we have chosen to cite capture the principal conclusions that can be gleaned from such reports, although others could have been chosen to portray similar results.

2.5 Other Background Factors, Including Attitudes, Motivations, Self-Perception, Aspirations, Intentions, and Expectations

There remains one last group of factors that has been found to influence students' achievement, after taking account of SES, individual ability, and the characteristics of schools. It is the attitudinal-motivational state of the student, broadly defined.

In the introduction, we stated our intention to restrict our discussion to summary results of studies of intermediate variables. The summary treatment is based upon two considerations:

1. The students' affective states are not background factors in the same sense as those presented earlier in this chapter -- they are intermediate factors in that they are influenced by earlier educational histories (among other influences), and that they affect later educational performance. Indeed, they might sometimes become the mechanism by which various factors have influence upon achievement; for example, school policies might be designed to improve the self-image of a disadvantaged group, leading in turn to improved student performance.

2. These affective states are subject to change, especially over a period of time, unlike the principal family characteristics and the demographic characteristics of the student.

In this chapter we consider affective states as background variables, even though they may be influenced by the educational process. In Chapter 3 we consider them as outcomes and study their association with more traditional kinds of background. This varied treatment is a reflection of the fact that attitudes, motivations, self-perceptions, expectations, and so on, are in fact *intermediate* variables rather than background or outcome variables.

2.5.1 Student-Parent Relationships

A number of studies have considered the relationship between parents and students as a background variable. Such relationships are usually expressed in terms of the student's perception of the interest his parents have in his school achievement, in their desire for a high level of educational attainment, their interest in the student's nonschool problems, and the assistance they give him with his homework.

Coleman (5) inquired concerning:

- Whether anyone read to the student before he started to school
- How good a student the mother (father) wants him to be
- How often students talk with parents about homework
- Father's (mother's) desire for child's educational level
- Frequency of parents' PTA attendance

Mayeske (32, pp. 373-4) tabulated the composite outcome measure by responses to these questions and the results for grade 3, 6 and 12 are shown in Table 39.

Table 39. Scale values assigned by Mayeske to various measures of parental interest in child -- Coleman data

	Grade		
Item	3	6	12
Whether anyone read to preschool child			
No	*	44.4	46.7
Once in a while	*	50.1	48.8
Many times, but not regularly	*	53.6	51.5
Many times and regularly	*	51.7	51.4
Don't remember	*	47.9	48.3
Nonresponse	*	40.0	43.9
Percent of variance explained		8.3	
Mother's desire for child's excellence			
One of the best students in class	50.5	51.0	51.7
Above the middle of the class	50.8	51.2	50.8
In the middle of the class	46.8	47.5	44.8
Just good enough to get by	48.3	44.2	41.2
Don't know	47.4	47.8	46.4
Nonresponse	43.5	40.5	46.7
Percent of variance explained	2.0	5.8	7.1
Father's desire for child's excellence			
One of the best students in class	50.7	51.1	51.7
Above the middle of the class	50.2	51.0	51.0
In the middle of the class	46.5	47.0	45.1

Table 39. Scale values assigned by Mayeske to various measures of parental interest in child -- Coleman data -- Continued

Item	Grade 3	 6	 12
Just good enough to get by	48.1	44.3	42.8
Don't know	48.6	48.2	46.8
Nonresponse	44.4	41.6	46.3
Percent of variance explained	2.2	5.6	6.7
School discussion with parents			
Just about every day		51.1	50.5
Once or twice a week		49.7	50.3
Occasionally, but not often		51.5	49.0
Never, or hardly ever		46.6	49.0
Nonresponse		42.2	45.0
Percent of variance explained		5.9	0.7
Father's desire for child's educational level			
Doesn't care if child finishes high school			45.4
Finish high school only			45.1
Technical, nursing, or business school after high school			47.3
Some college, but less than four years			47.2
Graduate from a four-year college			53.1
Professional or graduate school			55.7
Father not at home			47.4
Don't know			47.7
Nonresponse			45.2
Percent of variance explained			12.6
Mother's desire for child's educational level			
Doesn't care if child finishes high school			44.7
Finish high school only			45.0
Technical, nursing, or business school after high school			47.1
Some college, but less than four years			47.1

Table 39. Scale values assigned by Mayeske to various measures of parental interest in child -- Coleman data -- Continued

Item	Grade 3	Grade 6	Grade 12
Graduate from a four-year college			52.8
Professional or graduate school			55.0
Mother is not at home			47.8
Don't know			47.8
Nonresponse			44.5
Percent of variance explained			11.7
Frequency of parents' PTA attendance			
Not at all			49.8
Once in a while			50.0
About half the meetings			49.7
Most or all of the meetings			51.3
There isn't a parent association			53.6
Don't know			47.5
Nonresponse			44.3

* Coding of answers not clear.

Source: Mayeske (32, Appendix III)

Apparently, reading to the preschool child is associated with higher scores. Unfortunately, it is not possible to determine from the Mayeske data the correlation between this factor and, say, parental education.

Although there is an apparent association between parental desire for child's excellence and achievement, as well as between parental desire for the child to achieve a high educational attainment and achievement, one cannot help but wonder if these measures of parental attitude are influenced by performance. It would be surprising if they were entirely independent. In any case, both measures seem to be such that dichotomous response categories would suffice, i.e., parental desire for "above average" performance and for attainment of a four-year college degree. These are the points in the response categories where major changes in the criterion occur.

No remarkable differences appear in the association between outcome and the measure of parental attendance at PTA meetings except for the distinctly higher average for children who reported no PTA in the school. One can hardly believe that this is a causative relationship. It may be that schools with no PTA have a more highly selective student body.

The student questionnaire for the international study of literature education (Purves, 41) asked several questions concerning student-parent interaction.

10. How often does your mother or father help you with your homework? (three categories)

11. When you talk at home, do your parents

    always or almost always insist that you speak correctly?
    sometimes insist that you speak correctly?
    let you speak how you please?

12. When you show your parents anything you have written, do they

    always or almost always check your spelling?
    sometimes check your spelling?
    rarely or never check your spelling?

14. In your spare time at home, do your parents

    encourage you to read as much as possible?
    sometimes suggest you read?
    not mind if you never read?

15. When you get home from school, do your parents

    always or almost always want to know how you have done?
    sometimes ask about your school work
    hardly ever or never ask you about your homework?

Other questions concerned number of hours of TV-watching, hours spent reading, and so on.

Purves reported the multiple correlation coefficient (for the United States) of 0.39 between literature scores and the parental-interest variables, and showed that there was only a 0.002 increment in R^2 beyond the contributions of other variables (41, p. 417).

Thorndike (48, Table 6.1) used the same data as Purves, but with reading comprehension scores as the criterion. He created a "parental interest" composite from (1) parent's interest in school, (2) encouragement given the child to read, and (3) encouragement given the child to visit museums. He created a "parental help" composite from (1) whether they typically corrected his speaking, and (3) whether they typically corrected his writing. He found the following simple correlation coefficients for U.S. students:

	Parental Interest	Parental Help
Population I (10-year-olds)	0.08	-0.07
Population II (14-year-olds)	0.11	-0.20
Population IV (seniors)	0.16	-0.15

He reported (pp. 75-6), "In every country, it was found that the children who reported that their parents helped them with their homework tended to get poorer reading scores than those who reported that their parents seldom or never provided this type of help."

Bachman in Youth in Transition (1, Vol. II, pp. 17-21) studied interpersonal relationships of tenth-grade boys with their parents. Five indexes were constructed:

- Parental control
- Closeness to mother
- Closeness to father
- Parental consultation with son
- Parental punitiveness

The first did not show any relationship with the criterion variables, but the other four did. Since the treatment of such variables is more extensive than in most of the reports reviewed, Bachman's questions and scale values are reported, for reference, in Table 40. He described construction of the scales as follows:

> A total of 21 questionnaire items ..., were used to compute the measure of family relations. The total score on this scale consists of the mean of the scores for all available items, with up to five missing data cases allowed; in other words, a respondent had to provide answers to at least 16 of the 21 questions in order for a scale score to be computed. The scores for each response are indicated in parentheses; score values (ranging from 1 to 5) were assigned in such a way as to reduce distortion caused by missing data.

Bachman reported a small positive correlation between the family-relations index and socioeconomic level, small negative correlation with number of siblings, and positive correlation with intact homes. An indication of the amount of association with the Quick Test as a criterion variable is provided by Table 41. For this purpose Bachman's Eta of 0.16 may be interpreted as a simple correlation coefficient of that magnitude. Hence it is clear that the variable has small explanatory power alone and would have even less after account is taken of socioeconomic level, number of siblings, broken homes and possibly other variables.

Peaker, in a regression analysis of the Plowden data (8, Appendix 4) reported three parent-student variables:

- Aspiration for the child
- Literacy of home
- Parental interest in school work and progress

The variables are composites and the criterion variable is score in a test of reading comprehension. Analyses were done for three age groups, both between schools and among students within schools. Results of

Table 40. A composite measure of family relations -- Bachman

	% Answering,[a] and Score Value (in parentheses	
CLOSENESS TO FATHER		
When you were growing up, how did you feel about how much affection you got from your father (or male guardian)?		
Wanted and got enough affection	60	(4)
Wanted slightly more than I received	18	(3)
Wanted more than I received	14	(2)
Did not want affection from him	5	(1)
How often do you and your father (or male guardian) do things together that you both enjoy--things like playing sports, or going to sporting events, or working on things together?		
Several times a week	19	(5)
About once a week	29	(4)
Once or twice a month	21	(3)
Less than once a month	22	(2)
How close do you feel to your father (or male guardian)?		
Extremely close	30	(5)
Quite close	35	(4)
Fairly close	19	(3)
Not very close	8	(2)
How much do you want to be like your father (or male guardian) when you're an adult?		
Very much like him	26	(5)
Somewhat like him	36	(4)
A little like him	18	(3)
Not very much like him	8	(2)
Not at all like him	6	(1)

---

[a] Percentages do not add to 100 because missing data are not listed in this table. Missing data never exceeded 8%, and usually equalled 2% or 3%.

Table 40. A composite measure of family relations -- Bachman -- Continued

CLOSENESS TO MOTHER	% Answering, and Score Value (in parentheses)	
When you were growing up, how did you feel about how much affection you got from your mother (or female guardian)?		
Wanted and got enough affection	72	(4)
Wanted slightly more than I received	15	(3)
Wanted more than I received	7	(2)
Did not want affection from her	3	(1)
How close do you feel to your mother (or female guardian)?		
Extremely close	42	(5)
Quite close	37	(4)
Fairly close	15	(3)
Not very close	3	(2)
How much do you want to be like the kind of person your mother (or female guardian) is?		
Very much	26	(5)
Somewhat	38	(4)
A little	25	(3)
Not very much	9	(2)
Not at all	4	(1)
AMOUNT OF REASONING WITH SON		
How much influence do you feel you have in family decisions that affect you?		
A great deal of influence	19	(5)
Considerable influence	35	(4)
Moderate influence	26	(3)
Some influence	13	(2)
Little or no influence	6	(1)

Next we would like to get some idea of how often your parents (or guardians) do each of the following things:	Always	Often	Sometimes	Seldom	Never
Listen to your side of the argument.	18 (5)	30 (4)	32 (3)	13 (2)	5 (1)
Talk over important decisions with you.	12 (5)	32 (4)	34 (3)	14 (2)	5 (1)
Act fair and reasonable in what they ask of you.	19 (5)	36 (4)	33 (3)	9 (2)	2 (1)

Table 40. A composite measure of family relations -- Bachman -- Continued

PARENTAL PUNITIVENESS	% Answering, and Score Value (in parentheses)				
Next we would like to get some idea of how often your parents (or guardians) do each of the following things:	Always	Often	Sometimes	Seldom	Never
Completely ignore you after you've done something wrong	3 (1)	9 (2)	19 (3)	35 (4)	31 (5)
Act as if they don't care about you any more .	3 (1)	6 (2)	14 (3)	26 (4)	48 (5)
Disagree with each other when it comes to raising you	4 (1)	10 (2)	22 (3)	32 (4)	29 (5)
Actually slap you	2 (1)	7 (2)	19 (3)	31 (4)	39 (5)
Take away your privileges (TV, movies, dates).	3 (1)	9 (2)	25 (3)	35 (4)	26 (5)
Blame you or criticize you when you don't deserve it	3 (1)	13 (2)	32 (3)	36 (4)	15 (5)
Threaten to slap you	5 (1)	13 (2)	27 (3)	29 (4)	24 (5)
Yell, shout or scream at you	6 (1)	16 (2)	34 (3)	30 (4)	11 (5)
Disagree about punishing you	3 (1)	11 (2)	31 (3)	34 (4)	19 (5)
Nag at you	6 (1)	16 (2)	31 (3)	29 (4)	16 (5)

Table 41. Mean Quick Test scores for each category of family relations

Grand Mean = 108.5
Grand Standard Deviation = 12.5

Predictor Category	Weighted N	Percent	Unadjusted Mean	Standard Deviation
1 (poor)	106	4.2	108.78	13.03
2	199	7.9	106.30	14.61
3	344	13.7	105.69	12.65
4	482	19.2	107.88	12.21
5	555	22.1	108.83	13.25
6	427	17.0	110.93	11.31
7	259	10.3	110.92	10.87
8 (good)	90	3.6	109.60	9.03
9 Missing Data	52	2.1	102.32	9.67

Eta = 0.16

Source: Bachman (1 Vol. II, Table E-4-4)

the within-schools analysis are reported in Table 42. The percent of explained variance includes literacy which is perhaps approximately equivalent to education of parents in U.S. studies. Also, the analysis is a within-schools analysis while the other studies reported above are across all schools. Even with some downward adjustment for these factors the explained variance seems high in comparison with U.S. results.

On the basis of the results cited in this section it appears that relationships between parent and child do have an impact on school outcomes. However, substantially differing results have been obtained by different investigators. How much of these differences is due to the outcome measures, how much to the particular measures of parent-student relationships, and how much to the peculiar characteristics of the population being studied is unknown. Some additional perspective will be provided in the next section where joint relationships are examined in more detail.

Table 42. Variance in reading comprehension scores accounted for by parental variables -- Plowden data

Group and Variable	Simple Correlation	Percent of Variance
Top Junior Boys		
Aspirations for child	0.41	
Literacy of home	0.36	
Parental interest	0.42	26
Top Junior Girls		
Aspirations for child	0.57	
Literacy of home	0.37	
Parental interest	0.43	36
Lower Junior Boys		
Aspirations for child	0.27	
Literacy of home	0.34	
Parental interest	0.22	15
Lower Junior Girls		
Aspirations for child	0.25	
Literacy of home	0.26	
Parental interest	0.24	12

Source: Peaker (9, Tables 5.1 - 5.4)

2.5.2 Attitudes, Expectations and Kindred Variables

Bachman (1) used the Quick Test of Intelligence as the indicator of ability for most of his analyses and his study population was tenth-grade boys. However, we have used it as an indicator of achievement earlier in this chapter and continue to do so here.

Bachman reported correlations between the Quick Test and various affective measures as shown in Table 43.

Table 43. Correlations between various affective states and Quick Test scores reported by Bachman

Variable Correlated With Quick Test	Correlation Coefficient	Percent Explained
Self-concept of school ability	0.43	0.18
Positive school attitudes	0.07	0.01
Negative school attitudes	0.25	0.06
Need for social approval	0.15	0.02
Self-esteem	0.14	0.02
Social values	0.19	0.04
Ambitious job attitudes	0.28	0.08
Total internal control	0.22	0.05
College plans	0.30	0.09
Occupational aspirations	0.38	0.14

In Chapter 3 we discuss Bachman's findings with respect to the ability to predict these affective measures from other background factors.

Table 43 shows that self-concept of school ability explains 21 percent of the variation in Quick Test scores. One must presume, of course, that self-concept is nontrivially correlated with real ability. The next highest correlation is with occupational aspirations, which explains 14 percent of variation.

Three variables not shown in Table 43, because they are not affective states, motivations or aspirations, are religious preference, family political preference and political knowledge with correlation ratios with the Quick Test of 0.26, 0.13 and 0.36, respectively. They are reported here because they do not fit any other section of this chapter.

Mayeske (31) constructed an index called Family Process that included expectations for excellence, attitude toward life, educational plans and desires, and study habits. He was able to show that there was some correlation between achievement (as measured by his achievement composite) and these indices of student attitude. His results for third, sixth, ninth, and twelfth grades are shown in Table 44. The tendency for these measures to reach their maximum association with achievement in the ninth grade is worth noting.

Table 44. Simple correlations found by Mayeske between attitude indices and achievement -- Coleman data

	Grade			
Attitude Index	3	6	9	12
Expectations	0.17	0.26	0.39	0.35
Attitude toward life	0.13	0.38	0.47	0.42
Educational plans and desires	0.24	0.48	0.51	0.49
Study habits	0.34	0.37	0.36	0.23

Mayeske also showed that these attitude indexes were nontrivially correlated with his index of SES, the percent of explained variance in SES ranging for ninth grade from a low of 0.15 for Attitude Toward Life to a high of 0.29 for Educational Desires and Plans.

The international assessment studies have included a group of variables identified as "kindred attitudes and interests." For the literature evaluation (Purves, 41) this group of variables included:

- Age of students
- Expected occupation and education of the students
- Predominant sex of the students
- Amount of homework per week
- Amount of reading for pleasure
- Reading and viewing interests
- Literary-interest score
- Reading mystery and detective stories

Age and sex of students have been discussed elsewhere by us, and some of the items are specific to the literature test. Purves shows that these kindred variables (except for sex and age) account for six percent of the variation in senior student achievement in literature in the United States beyond the 24.9 percent accounted for by home background, age, sex, type of school, and instructional variables. They also account for 9.2 percent of the variance in 14-year-old scores beyond the 28.6 percent accounted for by the same other background variables. Thus, one must attribute nontrivial importance to these attitudinal variables in the prediction of literature scores.

Comber and Keeves (7) included in kindred variables science interests and attitudes and science reading in addition to expected education, expected occupation and hours reading for pleasure. They found that for high school seniors in the United States these kindred variables added four percent in explanatory power to the 35 percent accounted for by home background/school factors and learning conditions. Again, this is a nontrivial addition. If these kindred variables had been entered first in the regression equation they would have contributed substantially more than four percent. For example, the following simple correlation coefficients between the measures shown and science scores were reported by Comber and Keeves for twelfth-grade students in the United States:

Science attitudes	0.43
Expected education	0.32
Expected occupation	0.14
Hours reading for pleasure	0.18

The first, alone, would have accounted for over 18 percent. Of course, it is also correlated with a number of other background factors.

More will be said concerning intermediate-type variables in Chapter 3 where their analysis is considered both as an outcome and as an input to other outcomes. Here it is sufficient to note that their association with academic performance is nontrivial.

2.6 Summary

We close this chapter by presenting some results which show the proportion of variance in academic outcomes that can be accounted for by all background factors. Comber and Keeves (7) present the results shown in Table 45, where the outcome is total science test scores. The total percents accounted for range from 34 to 39 percent.

Table 45. Incremental percents of variance in individual student science test scores accounted for by all background variables studied, U.S. -- Comber and Keeves data

Variable or Composite	10-year-olds	14-year-olds	Seniors
Home circumstances	16	16	9
Age of student	-	1	1
Sex of student	1	4	8
Type of program	-	2	9
Type of school	1	-	-
Learning conditions	9	7	8
Kindred variables	7	6	4
Total	34	36	39

Source: Comber and Keeves (7, Chapter 9)

Purves (41) made similar estimates for literature scores. They are summarized in Table 46. Note that these percentages (61.5 and 53.0) are higher than the 34 to 39 percent in science, above. However, two other outcome variables, word knowledge and reading comprehension, have been included, and it seems reasonable to assume that they are relatively highly correlated with literature scores. Excluding them, the percents accounted for are 37 and 31 for 14-year-olds and high-school seniors, respectively, figures that are reasonably comparable with the Comber and Keeves findings.

Table 46. Incremental percents of variance in individual student literature test scores by all background variables studied, U.S. -- Purves data

Variable or Composite	14-year-olds	High School Seniors
Home background and reading resources, age, and sex	18.5	16.5*
Type of school/type of program	2.4	3.4
Instructional variables	7.7	5.0**
Kindred attitudes, interests, and other student attributes	9.2	6.0
Word knowledge	10.0	8.5
Reading comprehension	13.7	13.8
Total	61.5	53.0

* Includes number of siblings.

**Includes teacher and school characteristics.

Source: Purves (41, Chapter 6)

Bachman (1, Vol II, p. 72) reported 35 percent of variance in Quick Test scores accounted for by:

- Socioeconomic level
- Number of siblings
- Broken home
- Family relations
- Religious preference
- Family political preference
- Community size
- Race

Using only race, SEL, and number of siblings as predictors, he was able to account for 33 percent.

Mayeske, using the Coleman data (33, Chapter 6) found that he could account for about 53 percent of the variation in individual student achievement by family background and school measures. Attitudes, motivations and abilities would account for an additional increment, although we do not find the data summarized in that manner.

What emerges from this chapter, then, is that one might expect to account for from one-third to one-half of the variation in individual student measures of academic achievement by background factors that include measures of SES, family composition, student-parent relationships, race, sex, urbanism, geographic location and school and teacher variables. However, there are important differences in predictability by age groups, by racial-ethnic groups, and by outcome measures.

CHAPTER 3

OUTCOMES OTHER THAN ACADEMIC ACHIEVEMENT

3.1 Overview and Summary

The background and other factors that influence educational achievement are reviewed in Chapter 2, above, along with notes on measurement problems, methods of analysis, and the kinds of considerations required for a valid interpretation of the data. The present chapter explores the evidence on the kinds of factors that influence further education, occupation, income, job histories, and some other outcomes of the educational process.

Mushkin et al. (39) have identified a number of measures of educational outcome that are related to personal and societal objectives as follows:

Investment yield	(1) Advances in earnings with added schooling
	(2) Employment indexes (including unemployment rates by educational level)
	(3) Motivational behavior toward learning of children
Consumer effectiveness	(4) Patterns of consumption
	(5) Relative use of preventive medical and dental services
	(6) Reading expenditure patterns
	(7) Recreational program uses

Political effectiveness	(8) Voter participation and behavior
	(9) Group-interest participation (school board elections, community neighborhood groups)
	(10) Community services performance
	(11) Holding of public office
Personal effectiveness	(12) Capacity to reach out for satisfactory quality of life
	(13) Self-esteem
	(14) Internal-external control
	(15) Other attributes
Intellectual effectiveness	(16) Capacity for creativity
	(17) Knowledge about selected items
	(18) Completion of high school
	(19) Attendance at college or other post-secondary schools

The examined literature does not provide the basis for measuring associations between background and all of the above measures. However, Mushkin's list provides a valuable classification of outcomes for many purposes. It may be observed that most National Assessment exercises are measures of item (17) above, but some also fall under the heading of "political effectiveness" and "personal effectiveness." In this chapter most of the outcome measures are identified with items (1), (2), (13), (18) and (19).

In this study it will be seen that the various background factors have different patterns of influence on the educational outcomes reviewed in this chapter, compared with the outcomes reviewed in Chapter 2. But -- and this is an important finding in the present context -- the most influential factors are the same ones for both classes of outcomes. This result is a fortunate one from the viewpoint of research design. It directs attention and resources to relatively few factors -- and hence relatively few survey items -- which can be measured and used instead of attempting a broader and more superficial coverage, with increased costs, more serious response burdens, and attendant operational problems. Even the relatively complex sex-race interaction for early job experiences depends upon the same survey items found to be essential to the analysis of achievement scores.

Many different aspects of adult life can be regarded as educational outcomes in the sense that experiences connected with the processes of formal education can be influential. Some of these are expressive of intellectual development, such as eventual educational attainment, reading habits, choice of occupation, etc. Others are strongly economic: labor-force participation, employment history, periods of unemployment, income, welfare experiences, expenditure patterns. Others express societal roles: voting behavior (whether a person votes, rather than for whom), participation in community-betterment activities or volunteer work with those less fortunate. Important aspects of adult life also include familial roles and aesthetic life such as work in the graphic or performing arts or in creative writing.

All of the above kinds of characteristics of adult life fall within one or more of the 10 classes of NAEP objectives, presumably to be embodied in suitable exercises, especially in the young-adult age-group -- 25-34 years of age. Some of the instruments for these objectives are yet to be developed; those that have been designed are properly concealed from public view. In consequence, the discussion of outcomes other than educational achievement lacks the specificity of test items or survey items; rather, we deal with important areas which are nontrivially represented in the collection of NAEP objectives.

We also anticipate that there will be strong intercorrelations among the various items that purport to measure some one aspect of adult life: financial success, civic activity, intellectual life, humanistic or aesthetic experiences, etc. These interrelationships justify our proceeding without a review of specific instrumentation to be used by NAEP.

We first present some results of Duncan et al. (11), which explore the relation of background factors to educational attainment, occupation, and income for white male adults in Sections 3.2.2 and 3.2.3. Then the differences between black and white men are analyzed in terms of the same background factors and outcomes in Section 3.2.4. Using data from Project TALENT, attendance at college is studied for influences of SES, sex, and ability in Section 3.3.

Job experiences are also related to educational factors, sex, and race in Section 3.4. The Parnes studies (30, 40, 45, 46, 53) review labor force participation, unemployment and rates of pay and are discussed in Sections 3.4.2, 3.4.3 and 3.4.4.

Coleman (5, p. 275) points out that aspirations and motivations, especially those "toward further education and toward desirable occupations are partly a result of the home, and partly a result of the school. They play a special role, for they are in part an outcome of education, and in part a factor which propels the child toward further education and achievement."

Motivational and attitudinal influences on educational achievement were briefly treated in Chapter 2, Section 4. We turn to their role as outcomes in Section 3.5.

Section 3.6 summarizes the results of all of these investigations.

3.2 Relation of background factors to education, occupation, and income

3.2.1 Approach

For economy of exposition, we begin with an analysis of the relation of SES-type background factors -- father's education, father's occupation, and number of siblings -- to educational attainment, occupation, and income, along with the most important interactions among these three outcomes. By this device, we cover some essential ground with a single model, turning to other models and other studies for extension of results after the general form of the relationships has been set forth.

How do each of the three background factors affect educational attainment? Occupation? Income? And how does education affect occupation and income?

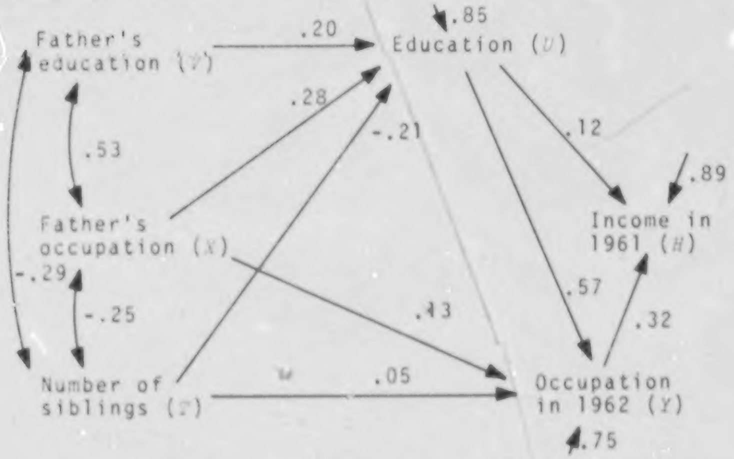
Duncan et al (11) investigate these questions with data from the March 1962 Current Population Survey (CPS) of the U.S. Bureau of the Census, which included a supplemental questionnaire, "Occupational Changes in a Generation," as well as the standard CPS items. Duncan and others developed the supplemental instrument. A sample of 20,700 respondents represented about 45,000,000 men in the U.S. civilian noninstitutional population between the ages of 20 and 64.

3.2.2 A Simple Model of Background Factors and Outcomes

For simplification, the initial model is restricted to non-black men with nonfarm background, 35-44 years of age, in the experienced labor force. (This imposes some problems on the interpretation of the data for NAEP purposes, to which we return in Section 3.2.3.) The model is extended to ethnic-group membership in Section 3.2.4.

The model is represented by Figure 2 from Duncan et al (11, p. 39). The figures on the lines are path coefficients. Path is not shown if the coefficient is less than 0.05 in absolute value.

Figure 2. Illustration of Duncan's Path Coefficients


.85
Father's education (V)
.20
Education (U)
.28
-.21
.12
.53
.89
Father's occupation (X)
Income in 1961 (H)
-.29
.57
-.25
.13
.32
Number of siblings (T)
.05
Occupation in 1962 (Y)
.75

The straight lines show significant influences, with the arrowheads indicating the presumed direction of influence. The curved lines with arrowheads on both ends represent associations with no such presumptions. The figures are path coefficients (or beta-coefficients). Arrows from outside the network represent uncorrelated residuals at the specified nodes. (A complete exposition of the model and its formal properties appears on pp. 18-30 of Duncan et al.).

Background factors (e.g., father's occupation) *directly* influence both the education and occupation of the respondent (his son, in this case). They also operate through *intervening variables*, affecting occupation through education and 1961 income through both education and occupation. (1961 data are the latest full-year incomes reportable in March 1962.)

3.2.3 Discussion of the Model

Several comments on this sample model are in order, in the context of the present study:

a. The model postulates *directions* of influence as well as *measures* of association in some cases.

These are plausible choices for the model displayed, with valuable information on direct and indirect influences of background factors. (Duncan states the ground rules impeccably in Chapter 1.)

b. The large contributions of residuals for intervening variables and outcomes (0.85, 0.75, and 0.89) demonstrate that the data fall far short of "explaining" variations in educational attainment, occupation, and income.

c. The substantially complete explanation of outcomes and intervening variables is *not* the principal objective of the NAEP project; hence the relatively small portion of variation associated with these background factors is not fatal. One of the objectives of NAEP is to construct *valid comparison groups* for the assessment of changes over time of scores on NAEP tests and exercises, protected from misleading and invalid comparisons because of changing composition (over time) of populations, especially in the dimensions of the background factors. Otherwise changes in scores for groups could occur from one period to another, even though no change occurred in the tested performance of any subpopulation.

d. The age interval (35-44 years) of the CPS data does not overlap the NAEP older-age group (26-35) in the above diagram. However, the simple correlations of the two age intervals are quite similar (page 38, referred to above), as are the partial regression coefficients (page 40). Duncan comments, "For illustration, the results are displayed in the graphic form for only one cohort, men 35-44 years old in 1962. Most of the important features of the results are, however, shared by all the cohorts." (p. 40)

e. The analysis is restricted to "experienced" labor force. It would therefore omit the relatively few men entering the civilian labor force for the first time -- probably mostly veterans. NAEP does not have a cohort for the years of heavy entrance of male full-time job-seekers -- 18-24 -- when the specific experience of this part of the adult life cycle would be most important.

f. The analysis excludes women. We present some results for women below.

g. The selection of a restricted part of the population for the analysis would have the effect of reducing the sources of variability, i.e., those associated with race, sex, very young adulthood, farm background. These attenuations in variability could lead to reductions in correlations and possibly the size of regression coefficients. In consequence, the importance of the influential variables is probably understated in the model, if one is interested in the population at large.

In the next section we review the results for an augmented list of background factors.

3.2.4 Comparisons of Outcomes for Blacks and Whites

The discussion immediately above was restricted to white men; we now introduce the ethnic-membership distinction, attempting to explain the same outcomes -- educational attainment, occupation, and income -- in terms of SES, number of siblings, and, additionally, race. Table 47 is taken directly from Duncan et al. - his Table 4.3.

The differences between Whites and Blacks are shown on the last line in original units -- siblings, years of schooling, occupational score, and dollars of income.

The sequence is so arranged that each column is estimated as subject to influences appearing to its left (family SES is taken as independent, and does not have its own column).

The several components of each difference are estimated by path-coefficient techniques. Black means for explanatory variables were inserted into White regression equations to estimate "equivalent" Black values.

The raw data show average years of schooling for Blacks of 9.4 years, for Whites 11.7 years, a difference of 2.3 years. One is interested in seeing how much of that difference is "due to" (in the

Table 47. Differences in means between White (W) and Negro (N) with respect to number of siblings, educational attainment, occupational status, and income, with components of differences generated by cumulative effects in a model of the socioeconomic life cycle, for native men, 25 to 64 years old, with nonfarm background and in the experienced civilian labor force: March 1962[a]

Number of siblings		Years of school completed		1962 Occupation score		1961 Income dollars		Component[b]
(W) 3.85		(W) 11.7		(W) 43.5		(W) 7,070		
	- .54		1.0		6.6		940	(A) [Family]
4.39		10.7		36.9		6,130		
	- .47		0.1		0.6		70	(B) [Siblings]
(N) 4.86		10.6		36.3		6,060		
			1.2		4.8		520	(C) [Education]
		(N) 9.4		31.5		5,540		
					11.8		830	(D) [Occupation]
				(N) 19.7		4,710		
							1,430	(E) [Income]
						(N) 3,280		
	-1.01		2.3		23.8		3,790	(T) [Total]

[a] Occupational changes from CPS.

[b] Difference due to: A, socioeconomic level of family of origin (head's education and occupation); B, number of siblings, net of family origin level; C, education, net of siblings and family origin level; D, occupation, net of education, siblings, and family origin level; T, total difference, (W) minus (N) = sum of components (A) through (E).

Source: Duncan (11) Table 4.3

arithmetical sense - not cause and effect) differences in SES and family size. The second column of Table 47 provides an answer. The interpretation of the second column is that 1.0 years of the difference (over 40 percent) is accounted for by the racial difference in SES (assuming that the White regression applies to Blacks); 0.1 years more by family size, and 1.2 (half) of the difference is due to other factors. The final component is the estimate of difference in educational attainment if the Blacks had the same SES and family size as Whites.

This result is for a particular population and date and for this particular model; a different collection of explanatory variables could change the relative importance of the components of the racial differences. The 1.2 years "unexplained" could be regarded as a racial difference, within the restrictions just stated, and without further identification or measurement of additional influential variables. This result warns of the need to "control" race in the comparison of NAEP scores for educational attainments from one year to another.

Occupation

The difference in occupation scale is 23.8 points, derived from white average of 43.5 and black average of 19.7. The scale is an occupational SES by Duncan (47) ("A Sociometric Index for All Occupations," in A. J. Reiss and others, *Occupations and Social Status*, New York, Free Press, 1961) from Census classification of occupations.

While family SES is important (6.6 points), about half, again, of the racial difference is allotted to other influences that SES, family size, and education.

Again, the treatment of occupation as an outcome of education (4.8, or 20 percent of the racial difference) and background factors required separate comparisons for at least these two ethnic classes in comparing NAEP scores for occupation over time.

Income

In this model (see also earlier Figure 2) income is regarded as being influenced by all four of the other factors. Education accounts for $520 of the total difference of $3,790, or about 14 percent. The residual for other factors is $1,430, about 38 percent of the total racial difference.

Again, ethnic groups should be kept separate in comparing NAEP scores for income, in order to avoid differences over time caused by changes in ethnic composition within any comparison groups.

3.3 SES. Ability, and Attendance at College

In the previous section, educational attainment was treated as an outcome; and the influences of background factors and other variables were described. A particular aspect of educational attainment is attendance at college. The following treatment illuminates two matters not included above: (a) the interactions of SES and ability, as related to college attendance; and (b) results for both males and females.

Tables 48 and 49 are taken from Project TALENT (14). They show consistent relationships between the probability of entering college and a combination of socioeconomic status and a measure of ability. There are consistent increases in probability of attending

Table 48. Probability of a male entering college by ability and socioeconomic index, Project TALENT.

		Socioeconomic Quarter			
Ability Quarter		Low 1	2	3	High 4
Low	1	0.06	0.12	0.13	0.26
	2	0.13	0.15	0.29	0.36
	3	0.25	0.34	0.45	0.65
High	4	0.48	0.70	0.73	0.87

college with increased ability and with socioeconomic level and also with increases in the combination. The combination of lowest SES and lowest ability yielded probabilities of attending college under 10 percent for both sexes, and high SES combined with high ability associated with probabilities of attending college between 80 and 90 percent. The patterns for the two sexes were quite similar.

Table 49. Probability of a female entering college by ability and socioeconomic index, Project TALENT.

Ability Quarter		Socioeconomic Quarter Low 1	2	3	High 4
Low	1	0.07	0.07	0.05	0.20
	2	0.08	0.09	0.20	0.33
	3	0.18	0.23	0.36	0.55
High	4	0.34	0.67	0.67	0.82

Note that increased ability has a much greater effect than increased SES for both men and women, as measured and scaled in this study.

The analysis did not include the dimension of ethnicity. It is probable that a substantial part of the effect of membership in ethnic groups is already accounted for in the SES measurement, since the SES index is a composite of nine measures including value of home, family income, father's occupation, and the education of father and mother, all of which will reflect some ethnic differences. This is not to say that no additional information will be provided by the addition of an ethnic dimension or that the differences between the sexes of ethnic minority groups would not be revealing.

Project SCOPE showed the unequal racial distribution of 9,735 entering college freshmen in California, Illinois, Massachusetts and North Carolina. Tillery

(49, p. 79) presents the following data on racial distribution of freshmen enrollments:

Level of Institution	Percent of Freshmen White	 Nonwhite
Less than 2 years	75	25
2 years, less than 4 years	87	13
4 years, B.A.	88	12
More than 4 years, M.A.	93	7
More than 4 years, Ph.D.	94	6

The higher percentage nonwhite in less-than-two-year institutions compared to institutions with higher level programs is evident.

Project SCOPE summarized other facets of college attendance, aspirations and changes in aspirations that are not summarized here, but which provide useful background for the researcher interested in this area.

3.4 Job Experience as an Outcome of Education

3.4.1 Approach

The greatest part of the literature on educational outcomes relates directly to academic achievement in some form. This may partly be explained by the relative ease of obtaining test results, grades, and related data. In one view of educational goals, the payoff in education should also show up in the early adult years in terms of job experience. It should not be presumed that other outcomes, especially taking care of children and households, are not important aspects of life which must be related to educational experience. However, a well-established goal of the educational system is preparation for useful employment -- either paid employment or voluntary work on some meaningful activity -- presumably related in some way to skills and abilities developed in the school years.

For all aspects of education with a strong vocational leaning, e.g., "commercial or secretarial" tracks or vocational subjects, the job experience is really a validation of the entire educational process. It is the real-life proof that learning and indeed grades on vocationally-oriented subjects do predispose a successful student to a successful work career.

The relative paucity of data might be understood most easily in terms of the cost and difficulties of obtaining quality data on both educational experiences and job histories for the same individuals. The most direct approach to this problem is the longitudinal survey.

For our purposes, the studies of Herbert Parnes at the Ohio State University Center for Human Resources Research (30, 40, 45, 46, 53) are the most relevant. Since the base-year survey was 1966, and there is a lag of several years between the initial survey and the reports which analyzed two and three-year followups, there are not yet many data for those in the first samples who have subsequently entered the labor market. Of the four groups of United States population sampled by Parnes, only the young women aged 14 to 24 would have had substantial numbers who were in school in the base survey year.

Other studies presumably will throw light on the subject. The Longitudinal Study of Educational Effects of the National Center for Educational Statistics conducted a base-year survey in the spring of 1972, with reports now well along in preparation but not yet available. The first-year followup is just now in preparation. The original sample was of high-school seniors, so the age group is quite well chosen for examining the transition from the educational system to the job market as those who take no further schooling seek jobs at once and the other groups gradually drift into the job market after their various kinds of post-secondary educational experiences. It will be several years before these data become available for the question at hand.

The American Council on Education sample of entering freshmen also has a longitudinal cohort which has great potential for associative studies. The data bank is described by Bayer, et al. (2) and examples of normative studies are Creager, et al. (8) and Bayer, et al. (3). One can hope that these excellent summary statistics will eventually be related to background factors.

In each subsection which follows, the data for males are presented, followed by corresponding information for females.

3.4.2 Labor-Force Participation

Turning to the Parnes studies, we first present information on the labor-force participation rates for males 14 to 24 years, who are not enrolled in school (see Table 50). The numbers shown on this table are population estimates, obviously, rather than samples. Hence the numbers of blacks who have completed elementary school are probably represented by quite small numbers in the sample. Not unexpectedly, the labor-force participation rates for boys 14 to 17 years of age who have not completed elementary school are low -- 74 percent for whites and 84 percent for blacks. By age 18 to 19, much of the gap between these dropouts and those with at least some high-school education has disappeared. The labor-force participation rates for the higher age groups for those who have not completed elementary school are not significantly different from those who dropped out some time during high school. Throughout the table, the labor-force participation rates for young white males tends to be higher than for young blacks, but by age 22 the differences practically disappear.

For the entire age group 14 to 24, those not entering high school have definitely lower participation rates, but they will not be evenly distributed through this age interval.

A substantial part of the explanation of nonparticipation in the labor force is traced to poor health or physical condition. For example, "white youths with some health problems have a participation rate of 90 percent compared with 97 percent for those with no such problems."

Table 50. Labor force participation rates, by highest year of school completed: Males 14-24 years of age not enrolled in school, by color - Parnes data

Highest year of school completed	Whites Total number (thousands)	Whites Labor force partici-pation rate, survey week	Blacks Total number (thousands)	Blacks Labor force partici-pation rate survey week
8 or less	699	90	227	92
9-11	1,188	97	323	91
12	2,573	97	346	96
13-15	607	98	46	100
16 or more	335	96	21	100
Total	5,402	96	963	94

Source: Parnes, et al. (40, p. 61)

A major explanation of labor-force participation of males 18 to 24 years of age is their marital status:

	Whites	Blacks
Married	100%	99%
Other	93%	92%
Average	96%	95%

Within this age group, there seems to be no difference traceable to the wife working or not working.

Parnes also conducted a brief analysis for the *influence of local unemployment rates* and found no particular differences in labor force participation except among blacks where the relationship is the reverse of that postulated by the "discouraged worker hypothesis." The black participation rate was 91 percent where 1960 unemployment was under 4.2 percent, and the participation rate was 96 percent where unemployment was 6.2 percent or more. Parnes calls this "inexplicable" but it might be traceable to the pressure on young men to seek jobs when other wage-earners in the household find themselves unemployed or working at less than their highest skill (and income), or working only part time.

Among women, educational attainment and the school curriculum are influential in their job experience. These matters are discussed under the section on rates of pay (Section 3.4.4) because of the close interrelationships.

3.4.3 Unemployment

We turn now to *unemployment rates of those actually in the labor force*. The data are shown in Table 51. "With some exceptions, unemployment rates of male youth tend to vary systematically according to school enrollment status, color, and age" (40, p. 65). Unemployment tends to be considerably higher for students than for non-students. (This might represent the quite restricted conditions under which full-time students can work.) Among both those in and out of school, the rate is higher for blacks than for whites. Unemployment also decreases with advancing age, but part of this is probably explained by the differences between students and nonstudents. In Table 51, looking only at those not enrolled in school, there is a sharp drop from age 16-17 to age 18-19 and a further drop -- much smaller -- from that group to the 20 to 21-year-olds. Above this age, the pattern becomes indistinct. Again, it should be noted that these very small percentages unemployed (in the upper age groups of the 14 to 24-year-olds) probably represent very few individuals in the sample, so much of the irregularity might be traceable to sampling variances for these very small cells.

There is a fairly strong relation between unemployment rates and highest year of school completed. (See Table 52.)

The unemployment rates for those who have not completed high school are much higher than for those who have completed high school but had no further

Table 51. Percent unemployment rates, by school enrollment status and age: males 14-24 years of age in the labor force, by color - Parnes data

School enrollment status and age	Whites	Blacks	Total
Enrolled in school			
14-15	14.9	17.3	15.2
16-17	12.9	23.9	14.3
18-19	13.4	21.2	13.8
20-21	4.0	3.2	4.2
22-24	3.7	2.8	3.6
Total 14-24	11.9	18.5	12.6
Not enrolled in school			
14-15	27.9	15.0	23.8
16-17	8.3	16.3	9.7
18-19	4.2	7.6	4.7
20-21	3.0	2.4	2.9
22-24	1.0	3.1	1.3
Total 14-24	3.0	5.5	3.4
Total age group			
14-15	15.4	17.1	15.6
16-17	11.7	21.4	13.1
18-19	8.4	10.3	8.6
20-21	3.3	2.5	3.2
22-24	1.4	3.3	1.6
Total 14-24	7.2	10.2	7.5

Source: Parnes, et al. (40, p. 66)

education. There are further declines for additional attainment through completion of high school (with the anomaly of the 13-15 grade group for whites). Overall, the blacks have higher unemployment rates than the whites.

Not only does school attainment relate to levels of unemployment, but the *high school curriculum also has an influence*. The levels for "general" education are high for both whites and blacks; the vocational

Table 52. Percent unemployment rates, by age and highest year of school completed: males 14-24 years of age in the labor force and not enrolled in school, by color - Parnes data

Highest year of school completed	Whites	Blacks
Total 14-24		
Less than 12	4.9	7.7
12	1.8	3.3
13-15	3.7	0.0
16 or more	0.0	0.0
Average	3.0	5.5

Source: Parnes, et al. (40, p. 74)

curricula are also high. The low unemployment rates are found among students who took commercial or college-preparatory courses (see Table 53).

Table 53. Percent unemployment rates, by high school curriculum: males 16-19 years of age in the labor force and not enrolled in school*, by color - Parnes data

High school curriculum	Whites	Blacks
Vocational	4.1	13.9
Commercial	0.0	0.0
College Preparatory	1.6	4.5
General	6.9	13.9
Total	4.9	11.8

* Includes only those respondents who have completed at least one year of high school.

Source: Parnes, et al. (40, p. 75)

It should be noted that Table 53 is not controlled for number of years in school. The population is restricted to those males in the sample who have completed at least one year of high school. Since the college-preparatory students have the lowest drop-out rate, it is possible that a number of vocational and general-education students have already dropped out of high school by the time the survey asked whether they were unemployed. The curriculum designations are therefore confounded with educational attainment, which was shown above to be related to unemployment rate.

3.4.4 Rates of Pay

We turn finally to another aspect of work history as an outcome of educational and other influences -- *levels of pay*. Levels of pay are related to educational attainment. This is not surprising in view of the rather close relation between education and occupation, and the superior pay of those occupations that go with higher educational attainment. Some details are given in Table 54.

Not surprisingly, the analysis shows that young men completing 12 or more years of school have higher hourly rates of pay than those who completed fewer years no matter whether one looks at craftsmen or operatives, or blacks and whites. Further, additional training outside of the regular school shows consistent increases in pay with the addition of such programs. The remaining striking feature of the table is the differential in hourly rates of pay between blacks and whites even after controlling for educational attainment and for classification as craftsmen or operatives. The control of occupation is not very precise, since there are many kinds of craftsmen (including foremen) and many kinds of operatives, but the differentials are large enough to be impressive even with the loose controls offered.

The rates of pay of *women* depend substantially upon their occupations; these, in turn, reflect educational attainment and school curriculum. The major racial differences between rates of pay arise through unequal access to favored occupations rather than in differences of pay within classes of occupations. These results are presented in Tables 55 and 56.

Table 54. Mean hourly rate of pay of craftsmen and operatives, by highest year of school completed and by extent of vocational training outside of regular school: employed male wage and salary workers 20-24 years of age not enrolled in school, by color - Parnes data

	Whites		Blacks	
Education and Training	Craftsmen, Foremen	Operatives	Craftsmen, Foremen	Operatives
Highest year of school completed				
11 or less	$2.47	$2.39	$1.47	$1.72
12 or more	2.76	2.73	2.23	2.19
Average	2.65	2.60	1.86	1.99
Extent of training outside regular school				
None	2.57	2.36	1.84	1.87
1 type of program	2.66	2.75	(*)	2.29
2 or more types of programs	2.78	2.94	(*)	(*)
Average	2.65	2.60	1.86	1.99

* Means not shown where sample cases number fewer than 30.

Source: Parnes, et al. (40, p. 98)

The first column of Table 55 shows the strong relationship between professional and technical occupations and educational attainment. Note the very high percentages, for both races, who have at least completed the equivalent of college. College graduation is rare for all other occupations, and for both races, with the exception of Blacks who are non-farm

Table 55. Highest year of school completed by respondents employed as wage and salary workers, by major occupation group and color* - Parnes data for women. (Percentage distribution)

Highest year of school completed	White collar				
	Professional technical	Nonfarm managers proprietors	Clerical	Sales	Total or average
	WHITES				
8 years or less	1	7	2	3	2
9-11 years	2	23	13	32	13
12 years	20	60	68	56	55
13-15 years	20	6	14	6	14
16 years or more	57	4	3	3	15
Total percent	100	100	100	100	100
Total (thousands)	863	235	2,305	315	3,718
Median**	16+	12.3	12.5	12.3	12.6
	BLACKS				
8 years or less	4	18	2	17	4
9-11 years	2	0	10	13	7
12 years	12	11	62	66	43
13-15 years	1	0	21	4	12
16 years or more	81	72	5	0	34
Total percent	100	100	100	100	100
Total (thousands)	118	7	167	26	318
Median**	16+	16+	12.6	12.3	12.9

* The tabulations from which the data in this table were compiled exclude respondents who did not report rate of pay. Consequently, universe totals in this table are somewhat smaller than the corresponding figures in other tables of the Parnes report.

** Median computed from grouped data.

Source: Shea, et al. (45, p. 106)

Table 55. Highest year of school completed by respondents employed as wage and salary workers, by major occupation group and color* - Parnes data for women. (Percentage distribution) -- Continued

Highest year of school completed	Blue collar	Domestic service	Nondomestic service	Farm	Total or average
			WHITES		
8 years or less	18	46	20	48	12
9-11 years	42	28	36	33	20
12 years	39	26	40	20	48
13-15 years	2	0	4	0	10
16 years or more	0	0	1	0	10
Total percent	100	100	100	100	100
Total (thousands)	1,266	83	766	41	5,875
Median**	11.3	9.4	11.5	9.2	12.4
			BLACKS		
8 years or less	30	47	32	72	23
9-11 years	32	46	40	28	31
12 years	34	8	23	0	30
13-15 years	4	0	4	0	5
16 years or more	0	0	1	0	11
Total percent	100	100	100	100	100
Total (thousands)	251	174	289	10	1,044
Median**	10.9	9.2	10.4	<8	11.6

* The tabulations from which the data in this table were compiled excluded respondents who did not report rate of pay. Consequently, universe totals in this table are somewhat smaller than the corresponding figures in other tables of the Parnes report.

** Median computed from grouped data.

Source: Shea, et al. (45, p. 106)

Table 56. Mean hourly rate of pay of respondents employed as wage and salary workers, by highest year of school completed, major occupational group and color - Parnes data for women

Major occupation group	8 years or less	9-11 years	12 years	13 years or more	Total or average
			WHITES		
White-collar	$1.88*	$2.05	$2.26	$2.79	$2.38
Professional, technical	1.92*	3.07*	2.71	2.97	2.91
Nonfarm managers, proprietors	2.28*	2.06*	2.47	3.41*	2.45
Clerical	1.74*	2.11	2.27	2.42	2.26
Sales	1.91*	1.64	1.63	2.99*	1.76
Blue-collar	1.78	1.99	1.94	2.02*	1.91
Domestic service	0.94*	1.06*	0.91*	---**	0.96*
Nondomestic service	1.40	1.40	1.63	5.57*	1.67
Farm	1.17*	1.80*	1.13*	---**	1.41*
Total or average	1.64	1.85	2.13	2.83	2.16
			BLACKS		
White-collar	2.15*	1.71*	2.20	3.20	2.63
Professional, technical	1.42*	1.73*	2.69*	3.55	3.34
Nonfarm managers, proprietors	1.90*	---**	0.69*	4.22*	3.24*
Clerical	4.10*	1.77*	2.23	2.33	2.26
Sales	1.65*	1.30*	1.73*	1.90*	1.61
Blue-collar	1.46	1.77	1.93	1.90*	1.79
Domestic service	0.81	0.99	0.91*	---**	0.89
Nondomestic service	1.95	1.41	1.49	1.78*	1.41
Farm	1.08*	0.77*	---**	---**	0.97
Total or average	1.19	1.44	1.92	3.06	1.78

* Mean is based on fewer than 20 sample cases.
** Dashes indicate empty cells.

Source: Shea et al. (45, p. 112)

managers and proprietors. If one searches the table, it is possible to find a good deal of heterogeneity throughout the pattern displayed, including the substantial variability of educational attainment of people in clerical and sales occupations. However, there is a fairly consistent pattern of those in domestic service and farm occupations having quite low educational attainment, again for both races. To a slightly lesser extent the same is also true of the other blue-collar occupations.

Hourly rates of pay are also related to *educational attainment* as shown in Table 56. It will be noted that many of the entries on this table are based on quite small cells in the sample. However, there are consistent relationships in the average for all White women and the average for all Black women, with the increasing hourly rate going up with years of education on most of the lines. The pattern of differences in hourly rates of pay, controlled for both occupational group and years of education, fails to show the consistent pattern that applied to Black men and White men. In general, the gain in hourly rates with increased schooling is somewhat greater for Black women than for White women. Parnes suspects some of the occupational classification in this analysis on the grounds that a number of "professionals" are in teaching, an occupation which generally commands a higher-than-average salary for professional-technical women.

Education curricula and subject specialization have repeatedly been shown to influence occupation. Among women, typing and sometimes shorthand are the most frequent paths to early employment, as shown in Table 57.

A review of the occupational distribution of white women shows a preponderance of clerical and sales jobs for a broad range of educational attainment, being exceeded by blue-collar jobs only for those who do not complete high school. Two-thirds of those with more than high school education have clerical and sales occupations, with most of the rest being professional and managerial. The second line of the table shows that those with typing and shorthand move substantially into these kinds of jobs. It should be noted that many of those without these

Table 57. Occupation of longest job between school and first marriage, by highest year of school completed, whether took typing and/or shorthand in high school, and color* - Parnes data for women. (Percentage distribution)

Occupation of first job	9-11 years			
	No typing or shorthand	Typing only	Typing and shorthand	Total or average
	WHITES			
Professional/ managerial	1	0	5	2
Clerical and sales	28	42	55	37
Blue-collar	48	32	29	40
Domestic service	1	3	3	2
Nondomestic service	22	22	9	19
Farm	0**	2	0	1
Total percent	100	100	100	100
Total number (thousands)	1,065	483	385	1,938
Horizontal percentage	55	25	20	100
	BLACKS			
Professional/ managerial	3	0	0	2
Clerical and sales	2	13	6	4
Blue-collar	22	22	11	21
Domestic service	29	19	10	27
Nondomestic service	36	45	73	40
Farm	8	2	0	6
Total percent	100	100	100	100
Total number (thousands)	293	50	22	366
Horizontal percentage	80	14	6	100

---

* Includes only respondents who attended high school and those who did not graduate from college.

** Percentage is 0.1 to 0.5.

Source: Shea, et al. (45, p. 145)

Table 57. Occupation of longest job between school and first marriage, by highest year of school completed, whether took typing and/or shorthand in high school, and color* - Parnes data for women. (Percentage distribution) -- Continued

Occupation of first job	12 years			
	No typing or shorthand	Typing only	Typing and shorthand	Total or average
	WHITES			
Professional/ managerial	9	10	5	7
Clerical and sales	63	66	80	73
Blue-collar	18	15	11	13
Domestic service	2	2	1	1
Nondomestic service	8	7	3	5
Farm	0	1	0	0**
Total percent	100	100	100	100
Total number (thousands)	846	1,900	3,451	6,234
Horizontal percentage	14	30	56	100
	BLACKS			
Professional/ managerial	10	1	7	6
Clerical and sales	27	28	55	33
Blue-collar	28	34	13	27
Domestic service	10	12	10	11
Nondomestic service	22	24	15	22
Farm	2	0	0	1
Total percent	100	100	100	100
Total number (thousands)	174	131	72	379
Horizontal percentage	46	35	19	100

* Includes only respondents who attended high school and those who did not graduate from college.

** Percentage is 0.1 to 0.5.

Source: Shea, et al. (45, p. 145)

Table 57. Occupation of longest job between school and first marriage, by highest year of school completed, whether took typing and/or shorthand in high school, and color* - Parnes data for women. (Percentage distribution) -- Continued

	13-15 years			
Occupation of first job	No typing or shorthand	Typing only	Typing and shorthand	Total or average
	WHITES			
Professional/ managerial	26	24	18	22
Clerical and sales	62	67	72	67
Blue-collar	8	8	4	6
Domestic service	0**	0**	0	0**
Nondomestic service	4	2	6	4
Farm	0	0	0	0
Total percent	100	100	100	100
Total number (thousands)	304	418	446	1,191
Horizontal percentage	26	36	38	100
	BLACKS			
Professional/ managerial	28	27	0	22
Clerical and sales	25	32	83	38
Blue-collar	11	30	0	14
Domestic service	11	5	8	9
Nondomestic service	26	5	9	17
Farm	0	0	0	0
Total percent	100	100	100	100
Total number (thousands)	42	20	15	77
Horizontal percentage	55	26	19	100

* Includes only respondents who attended high school and those who did not graduate from college.
** Percentage is 0.1 to 0.5.

Source: Shea, et al. (45, p. 145)

secretarial skills also move into clerical and sales jobs except for the less-than-high-school-education group, where a preponderance of those without secretarial skills become blue-collar workers or enter non-domestic service.

The pattern among Black women is quite different. Blue-collar, non-domestic, and domestic services dominate the occupational distribution even for those with typing, shorthand, or both. It is not until after graduation from high school that clerical and sales occupations become important, and then typically only for those possessing both typing and shorthand skills.

3.4.5 Summary of Job-Related Outcomes

In summary, it appears as though the job histories of both men and women, for Whites and for minority ethnic groups depend heavily upon education. To the extent that other variables are important, they seem to be the same kinds of background factors which this report has already cited as those most important for the associative explanation of educational achievement.

The general picture presented, then, is that the selection of background variables for NAEP to collect for the explanation of achievement scores also meets the requirement for explanatory factors for backgrounds other than achievement: educational attainment, attitudes important to employment success (see Section 3.5 following), and other aspects of adult life, employment and occupational aspects of job histories, as well as continuation of education for each level of education surveyed.

3.5 Attitudes and Motivations

3.5.1 Approach

In Section 2.5 we treated attitudes and motivations as background variables. Here we consider them as outcomes, recognizing that they are more properly classified as intervening variables. They are at once consequences of earlier educational experiences and other factors, and influences in later educational, occupational, and other behavior.

Attitudes toward school (Section 3.5.2) and work (Section 3.5.3), job satisfaction (Section 3.5.4) are reviewed, followed by educational and occupational aspirations (Section 3.5.5), and attitude toward government and society (Section 3.5.6).

3.5.2 Attitudes and Motivations Toward School

One central question on Coleman's twelfth-grade instrument (Question 59, p. 650) was: *"If something happened and you had to stop school now, how would you feel?"* The national distribution of responses follows:

	Percent
Like to quit	2
Don't care	4
Would be disappointed	12
Try hard to continue	36
Do anything to stay	45
(Nonresponse)	2
	100

The most common response, "Do anything to stay" showed 45 - 50 percent response for all five regions, and for metropolitan and nonmetropolitan communities alike. The White majority did not show higher attitudinal scores on this question than Black students for all region-community classes. Some other ethnic groups did score lower:

Mexican-Americans	37 percent
Puerto Ricans	35 percent
American Indians	36 percent

The lowest score ("like to quit") was small except for Puerto Ricans (15.9%).

Other attitudinal questions produced similar results. The Southern Region and Blacks reported more studying out of school than other groups (except Oriental-Americans) (Q. 61, p. 651). Blacks reported less staying away from school because they "didn't want to come" than others (again except Oriental-Americans) (263, p. 651). Again, the South reported a slightly lower rate of voluntary absenteeism.

3.5.3 Attitude Toward Work

Parnes investigated attitudes toward work, including job satisfaction, importance of money received, and other factors. Several of the questions are directed specifically toward women, and they deal with problems of child care during the work day, and similar items. Table 58 shows percentage differences in motivation to work at various occupations, by race. The data shows strong occupational differences. As anticipated, the professional and managerial jobs give "liking the work" far more emphasis than "good wages," these scores being identical for Blacks and Whites. At the other extreme, Blacks engaged in blue-collar occupations and domestic service indicated equal motivation for wages and job satisfaction.

It might be remarked that the finding that over 9/10ths of both Whites and Blacks reported liking their jobs either very much or fairly well might be regarded with some suspicion for some of the more menial jobs. Parnes cites other studies which display larger percentages of workers expressing dislike for their jobs. The group studied in this case is employed women 30 to 44 years of age.

Bachman (1) also investigates attitude toward jobs. There is only a moderate relationship between his values and attitudes and the corresponding items in the NAEP objectives for "career and occupational development." The Bachman "index of ambitious job attitudes" is a composite of "job that pays off" and "job that doesn't bug me;" the NAEP objectives probably have a number of components that would correlate well with this index. (Note the comment on non-availability of NAEP instruments in 3.1). The methodology and background variables are the same as

Table 58. Motivation to work of employed respondents, by occupation and color - Parnes data. (Percentage distribution)

Motivation to work	Professional, managerial	Clerical, sales	Blue-collar	Domestic service	Nondomestic service	Farm	Total or average
				WHITES			
Good wages	12	19	32	22	22	10	20
Liking the work	88	81	68	78	78	90	80
Total percent	100	100	100	100	100	100	100
Total number (thousands)	1,459	2,969	1,359	148	1,003	181	7,120
				BLACKS			
Good wages	12	27	49	50	38	85	39
Liking the work	88	73	51	50	62	15	61
Total percent	100	100	100	100	100	100	100
Total number (thousands)	147	215	264	269	323	34	1,253

Source: Shea, et al. (45, p. 180)

in the discussion of job information and achievement (grades) in Section 2.3. Bachman's data (Vol. II, p. 143) are as follows:

Bachman Variables	Percent of Variability Explained
Family relations (the largest single factor contribution)	11.2
Socioeconomic level (second largest)	5.5
All eight background factors	16.2
All eight background factors plus the Quick Test of Intelligence	19.5

3.5.4 Job Satisfaction

Parnes takes up the question of job satisfaction with men in Career Thresholds, (40). Again, there are differences in job satisfaction from white collar occupations to blue collar occupations. The differences become accentuated as the workers move from less than 12 grades attained (no difference between white-collar and blue-collar for Whites) to a fairly sharp difference for those with 13 grades or more -- 61 percent of the white-collar workers "like their job very much" contrasted with only 43 percent for blue-collar at the same educational attainment level. Among Blacks, there is an exception to this finding, in that 64 percent of the Black white-collar workers with less than 12 grades of attainment like their job very much.

Again, we find the summary picture similar for women workers. Occupational level is strongly influenced by educational attainment, and job satisfaction is related to occupation.

The relevance of these results to NAEP can be cited. The career and occupational development objectives include (#5) "Have positive attitudes toward work." This unit involves precisely some of the attitudes embodied in the Parnes questions -- recognizing the bases of various attitudes toward work, value of work in terms of societal goals, etc.

Since NAEP already collects information about education and occupations of those responding to the exercises, this does not suggest additional background variables necessary for the explanation of NAEP test results. Rather, it merely reaffirms the importance of those influences relating to high educational attainment, and probably to achievement.

3.5.5 Educational and Occupational Aspirations

Coleman (5, p. 275-281) discussed the educational aspirations and the occupational aspirations of his sample. In response to the question: "How far do you want to go in school?" the responses shown in Table 59 were given.

Table 59. Percentage distribution of replies of 12-grade pupils on "How far do you want to go in school?", for White and Black pupils in metropolitan and nonmetropolitan areas by region, and for selected minority groups for the United States, fall 1965 - Coleman data

Race and area	Want to finish college	Want to do professional or graduate work
White, nonmetropolitan:		
South	30	13
Southwest	39	20
North and West	28	14
White, metropolitan		
Northeast	32	21
Midwest	30	17
South	35	18
Southwest	35	19
West	33	22

Table 59. Percentage distribution of replies of 12-grade pupils on "How far do you want to go in school?", for White and Black pupils in metropolitan and nonmetropolitan areas by region, and for selected minority groups for the United States, fall 1965 - Coleman data -- Continued

Race and area	Want to finish college	Want to do professional or graduate work
Negro, nonmetropolitan:		
South	27	16
Southwest	30	22
North and West	19	14
Negro, metropolitan:		
Northeast	24	18
Midwest	27	19
South	29	22
Southwest	34	21
West	28	13
Mexican Americans, total	21	11
Puerto Ricans, total	18	14
American Indians, total	22	13
Oriental Americans, total	33	31
Other, total	20	17
Total, all races	30	17

Source: Coleman (5, p. 283)

For all races, 47 percent want to complete college, with some of these indicating even higher aspirations. When asked about *plans* to go to college, the answers are still strongly positive as shown in Table 60.

Table 60. Percent of 12th-grade pupils planning to go to college next year, for white and Black pupils in metropolitan and nonmetropolitan areas by region, and for selected minority groups for the United States, fall 1965 - Coleman data

Race and area	Definitely Yes	Probably Yes
White, nonmetropolitan:		
South	35	25
Southwest	50	23
North and West	35	26
White, metropolitan:		
Northeast	46	22
Midwest	37	25
South	41	26
Southwest	40	30
West	55	27
Negro, nonmetropolitan:		
South	30	38
Southwest	41	41
North and West	22	33
Negro, metropolitan:		
Northeast	31	31
Midwest	33	38
South	35	36
Southwest	43	34
West	48	37
Mexican Americans, total	26	34
Puerto Ricans, total	26	27
American Indians, total	27	35

Table 60. Percent of 12th-grade pupils planning to go to college next year, for white and Black pupils in metropolitan and non-metropolitan areas by region, and for selected minority groups for the United States, fall 1965 - Coleman data -- Continued

Race and area	Definitely Yes	Probably Yes
Oriental Americans, total	53	29
Other, total	32	29
Total, all races	38	27

Source: Coleman (5, p. 284)

Notice that the "definitely yes" answers are 38 percent, but the combination of "definitely yes" and "probably yes" are 65 percent, well in excess of the 47 percent stating that they wished to "at least complete college." As in other attitudinal questions the South and Southwest tend to state more positive attitudes than other parts of the country. And the White students tend to assign themselves higher scores than the Blacks. Also, consistent with the patterns above, some of the other minority ethnic groups score low (Oriental-Americans again being the exception). These generalizations apply both to stated wishes and to stated plans. Coleman points out that while fewer Blacks have definite plans to go to college, fewer also have definite plans not to attend. "This indicates the lesser concreteness in Blacks' aspirations, the greater hopes, but lesser plans" (p. 279).

These data do not allow for explanations beyond ethnic groupings and regions of the country. In the section that follows, Bachman (1) relates occupational aspirations and college plans to the Quick Test and background factors, but he has no educational inputs or explanatory variables as such. After summarizing these results, we turn again to the Parnes studies where young adults' occupational aspirations are related to a number of educational characteristics (but not to achievement scores).

Bachman's index of self-concept of school ability was formed from answers to three questions (1, Vol. II, p. 92).

- How do you rate yourself in school ability compared with those in your grade in school?
- How intelligent do you think you are, compared with other boys your age?
- How good a reader do you think you are, compared with other boys your age?

He found intelligence (as measured by the Quick Test) to be the strongest predictor of self-concept, accounting for 21 percent of the variation in the self-concept index. He also found that he could account for 16 percent of the variation in self-concept by eight background variables (indices) whose simple correlations (as measured by Eta) with self-concept are listed below:

Socioeconomic level	0.33
Number of siblings	0.21
Broken home	0.07
Family relations	0.19
Religious preference	0.18
Family political preference	0.13
Community size	0.11
Race	0.06

Note again the dominance of socioeconomic level which alone accounts for 11 (0.33 squared) of the 16 percentage points of variation explained by all eight background variables.

He also found that combining scores on the Quick Test with scores on the eight background factors permitted him to account for 29 percent of total variations in self-concept of school ability. Note that cause and effect are not to be implied from these data, since one could easily argue that self-concept is one of the determinants of achievement and the Quick Test has a strong relationship to achievement.

Bachman also investigated a number of other affective measures. Table 61 summarizes some of his principal results. Family relations, an index of 21 items dealing with relationships between respondents and their parents, emerges as the dominant predictor for seven of the nine affective variables shown in Table 61. For these seven outcomes, socioeconomic level alone accounts for no more than five percent of the variation. Thus, it appears that if outcomes of the kinds listed in Table 61 are to be adjusted for background, other variables than SES will have to be used. It could easily be argued that family relations is not a pure background variable since it can be influenced to some extent by education and social action programs. The Bachman study shows convincing evidence of its importance, however.

In the opening year of his five-year longitudinal study, Parnes investigated the educational aspirations of males 14 to 17 years of age, as shown in Table 62. Parnes comments that the general pattern is unrealistic in view of its implications for the distribution of occupation in the economy and he predicts that the level will drift down in the followup years of the longitudinal survey. However, in the first followup survey (1967) this pattern really does not yet emerge. The whites lowered their educational objectives 15 percent, while the blacks lowered objectives by 20 percent. However, approximately 14 percent of both color groups raised their sights for additional education, yielding trivial net declines. In general, the highest occupational and educational aspirations "are associated with urban rather than rural residence, high socioeconomic status of family, enrollment in college-preparatory curriculum, positive attitudes toward school, spending above-average amounts of time on homework, and having above-average knowledge of the world of work" (40, p. 185).

3.5.6 Attitude Toward Government and Society

As in attitudes toward jobs, Bachman has a somewhat different approach with respect to attitudes toward society than seem to be implied by the NAEP Citizenship Objectives. He scales "political knowledge" by identification of student's senators and a

Table 61. Associations observed by Bachman between background factors and various attitudes, self-esteem, affective states, values and attitudes.

Outcome measure	Percent of variance accounted for			Background variable with highest association	
	By background	By background and quick test	By SES	Variables	Percent accounted for
Positive school attitudes	0.15	0.15	0.01	Family relations	0.12
Negative school attitudes	0.18	0.20	0.04	Family relations	0.15
Need for social approval	0.13	0.14	0.02	Family relations	0.09
Self-esteem	0.16	0.17	0.02	Family relations	0.13
Social values	0.14	0.15	0.02	Family relations	0.13
Ambitious job attitudes	0.16	0.18	0.06	Family relations	0.11
Total internal control	0.11	0.13	0.03	Family relations	0.09
Political knowledge	0.13	0.20	0.08	Socioeconomic level	0.08
College plans	0.13	0.16	0.09	Socioeconomic level	0.09

Source: Bachman (1, Vol. II)

Table 62. Educational aspirations: males 14-17 years of age enrolled in elementary or high school, by color - Parnes data

Number of years of education desired	Percentage distribution		
	WHITES	BLACKS	TOTAL
12 or less	27	36	28
14	12	14	12
16	44	42	43
More than 16	18	9	16
Total percent	100	100	100
Total number (thousands)	5,298	827	6,125

Source: Parnes, et al. (40, p. 174)

few other high officials (1, Vol. II, Table 8-9 on page 156), whereas the NAEP citizenship objectives relate principally to the structures and functions of governments and to the citizen's participation in civic life. The explained variances in political knowledge is low in any case: 15 percent of the total variability for the eight background variables and 21.7 for these plus Quick Test scores. (1, Vol. II, Table 8-10 on page 157.)

There are other overlaps: NAEP's concern for welfare and dignity of others, rights and freedoms of all individuals, maintaining law and order (all from the citizenship objectives) might relate to Bachman's *social values* whose six scales comprise social responsibility, social skills, honesty, kindness, self-control, and reciprocity. The only consequential background variable -- family relations -- associates with 12.5 percent of the total variance of the social values composite scale; the eight background factors explain 15.1 percent and these plus Quick Test account for only 16.8 percent, i.e., leaving almost 85 percent of the variation unaccounted for (1, Vol II, Table 8-1 on page 140).

3.6 Summary of Findings for Outcomes Other Than Achievement

NAEP objectives cover a variety of behaviors that could be regarded as educational outcomes. In this chapter we reviewed some studies that investigated the relationships of background factors to some of these outcomes.

SES -- parent's education and occupation, and family size -- proved to be important enough in the determination of eventual educational attainment, occupation, and income that comparisons of NAEP scores would be fostered by control of these sources of variability. Membership in ethnic group also is associated with differences in outcomes, and also should be controlled in order to avoid misreading changes in ethnic composition as valid trends or differences in NAEP scores.

Similar results were also found for SES and attendance at college; but here a measurement of student ability also played a major role.

Job experiences also were conditioned by SES, especially by occupation of the head of the household. It was in occupation and income that major racial differences appeared, and also the interaction between race and sex; the jobs evidently available to black women showed a substantial wage disadvantage, even when controlled for parent's occupation, years of education, and curriculum.

Overall, the portion of total variability explained is fairly small, indicating the presence of important influences not included in the model or the analysis. However, background factors are too influential to neglect in making comparisons of NAEP scores.

CHAPTER 4

CONCLUSIONS

The principal large studies investigated show that one can expect to account for between 20 and 50 percent of the variation in academic outcomes (as measured by test scores) by variation in sex, race, home and family background, school characteristics and motivations, expectations, attitudes, and desires of individual students. Within this range there is substantial variation in association between background variables and outcomes, depending upon the nature of the outcome, the age group of the students, and the specific background factors considered. For example, sex is a more important predictor of scores in twelfth-grade science than in science scores for ten-year-olds, or (let's say) scores in reading at any age.

Other outcomes such as college attendance, occupational goals and achievement, self-esteem, educational attainment, and societal attitudes are generally related to the same kinds of background variables. For example, socioeconomic status and ability are primary determinants of college attendance, but the relationship of the child to his family is an important determinant of attitudes toward school, self-esteem and social attitudes.

4.1 Measures of SES

Almost every major study has included some measure of socioeconomic status (SES). One can expect SES alone to account for between 10 and 25 percent of the variation in academic scores. Most common measures of SES include occupation of parents, education of parents, and items in the home.

Occupation of father (or if no father in the home, of the mother) is usually obtained from teachers or other school personnel for children in the lower grades and from the students themselves in the upper grades and high school. Sometimes the response is simply a written reply to the question, *"What does your (the pupil's) father do?"* Answers are subsequently coded into from three to over 100 classes.

The International Assessments have used nine classes with codes that are approximately linearly related (by presumption) with income. Bachman (1) and others have used Duncan's scale (47) which contains over 100 categories.

The other common mode of recording occupation is to provide a precoded scale and to permit the pupil (or teacher) to select the category most nearly like the father's occupation. Coleman (5) used ten categories for this purpose. There is insufficient evidence in the literature to permit one to choose between the two methods of recording occupation on the grounds of accuracy. However, the precoding of occupation by the respondent himself is substantially less costly and appears to give satisfactory results.

While frequently only the occupation of the father is requested (replacing it by occupation of the mother if there is no father) it is common practice to request educational attainment of both father and mother. When one is not present the educational attainment of the other is used for both. Again, the response is usually given by the teacher for younger grades and by the pupil himself for upper grades.

Educational attainment is sometimes recorded into precoded classes such as: less than eighth grade, eighth grade but no high school, some high school but did not graduate, high school graduate but no college, some college but did not graduate, college graduate but no postgraduate work, and some graduate school. In other cases, such as the International Assessments, attainment has been recorded in terms of number of years of formal schooling. There is little evidence of difference in predictability and the former method seems more natural for U.S. students. Whether there are enough students whose parents were educated in a foreign country to make years of schooling a preferred method of recording is a factor to be considered.

A great variety of items in the home has been considered by various researchers. However, they tend to be in one or more of three classes:

- Those that represent economic status (e.g., automobile, telephone, TV, etc.)
- Those that represent an environment conducive to educational attainment (e.g., books in the home, daily newspaper, encyclopedia, dictionary, etc.)
- Those that represent a family interest in a specific subject matter such as art, music, literature, or science (e.g., musical instruments, art and literary magazines, paintings, scientific journals, etc.).

The first two categories are of interest to all educational outcomes, and the third category is tailored to specific educational outcomes. This pattern could easily be adapted to NAEP usage because of the grouping of exercises according to educational objectives.

THere are strong intercorrelations among the items in each of the first two groups, but smaller correlations between items in the first group and items in the second group. The inference is that there is a lot of freedom of choice among items to represent each class, but that both classes should be represented. Within each class a few items (as long as they are in any sense reasonable choices) will account for a significant part of the variance but adding items will produce smaller and smaller increments of explained variance so that relatively few (less than a half-dozen) items in each category will explain essentially all of the variance explainable by that category of item.

A study by RTI (43) investigated the response errors inherent in asking pupils about the occupations of their parents, their educational attainment, and the items in their homes. Pupil responses were compared with parental responses. When three categories of occupation and three categories of educational attainment were used it was found that percent consistency ranged from 52 percent for the lowest class of father's occupation to 85 percent for the lowest class of mother's occupation, and from 41 percent for fathers who did not graduate from high school to 76 percent for mothers who are college graduates. The results are for eighth-grade students. Results were somewhat poorer for third grade. Surprisingly, consistency between parents and pupils was lower for educational attainment than for occupational status.

However, consistency was relatively high for questions concerning items in the home, being 90 percent or better for television set, telephone, record player/hi-fi/stereo, refrigerator, automobile and vacuum cleaner for both third and eighth grade. Certain items have more discriminating power than others, however, by virtue of the fact that they are owned by a smaller proportion of the population. Encyclopedia, daily newspaper, magazine subscriptions, vacuum cleaner, and number of books tend to be of this nature. The RTI results are worth careful examination prior to implementation of any extensive collection of SES indicators.

4.2 Sex

Sex has differential explanatory power depending upon the age group and outcomes studies. There appears to be no reason why sex should not continue to be recorded.

4.3 Race

Race is highly correlated with measures of SES. However, the principal studies have shown that there is an additional contribution to the explanation of variance beyond that which can be attributed to SES. This is particularly true among young adult job experiences, and especially so for women. There is evidence of substantial differences among the non-Black ethnic minorities such as Mexican-Americans, American Indians, Puerto Ricans and Orientals. Even though representation of some of these is small in the NAEP sample, there appears to be sufficient evidence for consideration of separate recording of them.

4.4 Ability

Traditionally, ability has been considered a major predictor of educational outcomes. The major studies we have examined show that much of the ability measure is accounted for by SES and other background factors. There still remains some additional contribution of ability, however. Difficult problems of measurement may make it impractical to consider ability as a background variable for the National Assessment application.

4.5 School Factors

The Coleman study (5) was aimed primarily at observing differences among schools and relating those differences to educational outcomes. The literature is full of evidence that school differences, as measured by average SES of students, racial mix of students, and various characteristics of the school and the teaching staff, are important in explaining differences in education outcomes. However, most of those differences can be accounted for by differences in student background characteristics because of the high correlation between school characteristics and pupil characteristics. This had led some authors to conclude that schools do not make a difference, but one perhaps could also argue that, given the characteristics of the school, the backgrounds of the students do not make a difference. The point is that student background and school characteristics are interlocked and together they account for a substantial portion of variance in student outcomes.

4.6 Motivation. Attitude, Expectation and Other Measures

A number of intermediate variables (i.e., intermediate between the fixed background of the pupil and his educational achievement) have been shown to have an effect on educational outcomes. The interpretation of these measures as background variables is subject to the criticism that they, themselves, are at least partially an outcome of the educational process, and that all are nontrivially correlated with SES and ability. If they are to be considered in the educational model the evidence is that family relations, self-concept of school ability, ambitions, job attitudes, and plans for further education are important variables. The method of path coefficients used by Duncan et al. (47) Hauser, (21) and others provides a particularly rich methodology for the analysis of intermediate variables.

CHAPTER 5

EPILOGUE

5.1 Introduction

One of the suggestions that came from the discussion of the initial draft of this report with a review panel (see Preface for membership) was that we prepare an epilogue to the report. The epilogue would be a brief summary of impressions and recommendations that could go beyond what we felt free to do in the main text of the report, i.e., the simple summarization of the influence of background variables as seen in the literature. We have also developed some perspective on the principal studies that have been used to show associations between background and educational outcomes, and we would like to present some generalizations about such studies and the need for further study. Finally, the process of measuring background variables, and of measuring association, requires some further comment. These considerations, together with the suggestions from the review panel, have motivated us to write this chapter, which is not necessarily encompassed within the scope of the contract but which we feel may permit the reader to benefit more fully from our experience in accomplishing the literature review.

Throughout this report the views expressed are those of the authors and do not necessarily represent those of NAEP. This caution is particularly relevant to this chapter.

5.2 Methodology for Estimating Association

Percent of explained variance was chosen as the principal measure of association in this study partly because it appears so frequently in the literature. It is also a particularly effective measure since it is scale-free and relatively easy to compute.

It is by no means necessary to use linear regression in the computation of percent of variance explained by background factors, although the technique has been widely used. One could form mutually exclusive and comprehensive population subgroups, as illustrated by:

> Black females, age 13, living in urban areas in the Northeast, with parents in unskilled occupations and with less than high school education, etc.

Any tested child would fall into one and only one such subgroup, and the boundaries of the subgroups could be set so that some minimum number of sample cases was in each. Then the estimated percent of variation accounted for could be estimated and tested for significance by the usual methods of analysis of variance. In some ways, this kind of analysis has much to recommend it since it makes no assumption about relationships among the background variables. It is feasible in this framework to identify population subgroups whose performance and progress can be measured over time.

The analysis is not satisfying to some, however, because it does not permit one to say how much of the variation is "due to" education of parents, or occupation of parents, or type of community, etc. Regression analysis *appears* to do just that, but the appearance is deceiving. If one uses a single background variable he may "account for" (say) 10 percent of the variance in the criterion. Another variable may account for 8 percent and the two together (say) 14 percent. Apparently, the first adds 6 percentage points to the explanation by the second and the second adds 4 percentages points to the explanation by the first. But the situation may change substantially as soon as another variable is added. Mayeske and others have made useful attempts to identify unique contributions to variance of a set of variables. Nevertheless, as they fully recognize, what appear to be unique contributions of each variable are only unique with respect to the particular subset chosen and with respect to the linearity and additivity assumptions inherent in the regression model. Addition of variables reduces the so-called unique contributions and may increasingly call into question any assumption that the effects of the background variables are additive.

Even though illusory, the assignment of explanatory power to certain background variables may have considerable merit in serving NAEP's analytical goals. However, after the choice of background variables has been made there may be advantages in making comparisons group by group rather than in making comparisons of adjusted outcomes. We believe this less sophisticated approach to analysis is more enlightening to the extent that it is feasible. However, the sizes of the groups become a problem, and one may then need to resort to adjustment in an effort to make average or over-all comparisons in which some of the effects of associated variables have been removed.

Statistical adjustment is frequently undertaken to answer such questions as "How much (if any) would outcomes in the South be different from the national average if the South had the same average SES as the entire nation, the same racial distribution, the same schools, the same family size and composition, and so on?" The answer to the question is (hopefully) a measure of "the Southern effect," whatever that means. Although much of the literature warns against interpretation of differences in background as "causes" of differences in outcomes (and we have repeatedly done so in this report) there is a tendency to do so and, in this example, to say (incorrectly) that residence in the South "causes" the adjusted difference from the national average.

Quite aside from the matter of improper attribution of cause is the problem of the measurement itself. Adjustment requires assumptions about a model and, for computational convenience, the model that usually is assumed is one of a linear combination of separate background effects. There is abundant evidence in the literature we have studied that such an additive linear model is not a good one and that, in fact, almost every background factor is correlated with every other background factor so that, at least, the most important interactions should be built into the adjustment model. This could easily be done with modern computing machinery.

In any case, the observed differences after adjustment are presumably more nearly comparable than they were before. That is, the adjustment process will partially correct for factors which otherwise would bias the comparison. The amount of the correction is highly dependent upon the validity of the model, but even an ill-fitting model may make the data more comparable than they were prior to adjustment.

The apparent small magnitudes of explained variance reported in the study may be of some concern to the reader. After all, if 20 percent of the variation in the criterion is accounted for by a given set of variables then 80 percent must be unaccounted for, and that seems to be the dominant fact. One must remember that prediction is not the objective -- improved comparability is. Accounting for 20 percent of the total variation by classification (as above) or by regression or other adjustment procedures can add substantially to the interpretation of comparisons.

Few of the studies we have examined have paid any attention to the errors and especially the biases in estimation of the percent of explained variance. The California studies (4) did. The least-squares method of fitting regressions assures a minimum residual sum of squares (by definition) and hence a maximum of explained variation *for the set of sample data to which it is applied*. It will not fit another sample from the same population so well, hence it overstates the variation explained by the "independent" variables. The overstatement increases as the sample size decreases and as the number of variables increases since the fitted regression, under these circumstances, becomes more and more a function of the particular sample values used.

A simple way of testing the validity of regressions is to separate the sample into random halves, computing the regressions for each half independently and comparing the results. If some stability is evident, then the combined data can be used for final estimates. The exercise is particularly relevant when one is attempting to choose the most important among a set of candidate background variables, and in estimating the proportion of explained variance.

5.3 Sources of Errors

Within the resource limitations of this study it has not been possible to discuss the possible impact of measurement errors on the results of each study. In fact, most of the studies have given little attention to the matter.

Errors in surveys are of two kinds: (1) sampling errors or errors in coverage and (2) errors in measurement. Errors in coverage can be controlled within any prespecified limits by increasing the sample size, provided that the sampling procedures are unbiased and that responses are obtained from all sampling units. The first condition is frequently met and the second almost never.

As the rate of nonresponse increases one worries more and more about the selectivity of those who do respond. For example, students who are absent on the day of testing may tend, more than average, to lack motivation, be from low-SES families, be in particular racial/ethnic groups, and so on and, what is of more concern, their knowledge that testing is to be done may influence their attendance adversely, particularly if their performance is likely to be poor.

National Assessment emphasizes the importance of response rates, and that emphasis adds substantially to the quality of the assessment results. Most of the studies we have examined, however, are not explicit with respect to the procedures used to obtain response, although in at least half of them overall response rates are quoted. In general, it has not been possible for us to assess the impact of nonresponse on the results cited, but the consistency in results among the major studies, where a variety of survey procedures was used, lends support to the assumption that nonresponse may have had only a modest impact on the principal results cited.

The extent of errors in measurement, e.g., students marking their papers incorrectly when they in fact know the answer to a question, incorrect reporting of parental education and occupation, incorrect reporting of items in the home, etc., are not adequately known and few of the studies have attempted any kind of validity assessment. An exception is the NAEP-sponsored study to examine the accuracy with which students report items in the home, occupation and education of parents.

It is a peculiarity of judging degree of association by percent of variance explained that large errors in relatively few observations can have a major impact on the estimates. This is particularly true when many "independent" variables are used or separate regressions are used for small classes.

We know that errors of measurement decrease the amount of explained variance, perhaps substantially, but we have little evidence as to how much of such underestimation may have occurred in the major studies we examined. Also, adjustments designed to increase comparability of groups are likely to be underadjustments if there are measurement errors in the independent variables. Therefore the adjustments may leave differences unexplained that could have been explained by more accurate measurement.

5.4 Limitations in the Interpretation of Relationships from Survey Data

Throughout this report we have warned against inferring cause and effect from associations among variables, regardless of the strength of those associations. Inability to assign cause and effect is inherent in all analyses of survey data where there has been no opportunity to randomize the assignment of persons to treatment subgroups for which outcomes are to be measured.

This limitation is of little consequence when outcome data are adjusted for fixed background characteristics (e.g., race, sex, family SES, etc.) in order to make subgroups more comparable with respect to the variables used in adjustment. It may be of major consequence, however, if cause and effect are attributed to association with intermediate variables and educational policy is set so as to improve outcomes by improving the intermediate variables.

Nonexperimental data, such as have been presented in all of the studies we have reviewed, are particularly ineffective in testing hypotheses about cause. Frankel (16) has provided some evidence of the applicability of statistical distribution theory to tests of hypotheses in survey data, but the real issues are concerned with the possible validity of alternate causal hypotheses. The simple hypothesis that Black children perform as well as White children on standard reading tests can be quickly rejected. Where the difficulties arise is in hypotheses concerning the source or cause of the difference: Cultural or racial bias of the tests? Educational disadvantage of the home environment? Quality of schools and instruction? Differences in ability to learn? Differences in language used in the home and on the street? Economic deprivation of the family? Number of siblings? Heredity or environment (nature or nurture)? Recent interpretations of evidence on school effects, or lack of them, involve serious interpretation risks of this sort.

Acceptance or rejection of any of these hypotheses requires assumptions about interrelationships among the variables which, themselves, are incapable of test without assumptions about other relationships. Thus, the process is circular unless there has been randomized assignment of subjects to educational treatments.

Fortunately, the unscrambling of such hypotheses, or the adjustment of data to increase comparability on certain independent variables, is not necessary to setting important aspects of educational policy -- it would be helpful, of course. That certain subgroups of the population lag behind their peers in performance on measures of achievement is a fact, and presumably NAEP's functions include the measurement of that gap and its monitoring over time. The extent of the gap may be used as an indicator of the need for funding or for focus on certain subgroups in setting educational policy. Proof of the effectiveness of the tactics of education, that is, the actual methods for bringing lagging subgroups up to a higher level of performance, cannot be expected to come from the examination of survey data. Such data can be helpful, but may not be sufficient.

5.5 Unit of Observation and Analysis

The National Assessment administers relatively few exercises in any one age group in any one school, and for this and other reasons school performance measures are never published. Thus, the interest of NAEP is entirely in the student, not the school, as the unit of analysis for the dependent or outcome variables. However, there may be interest in school characteristics as independent or background variables that may account for some of the variability of the dependent or outcome variables.

An important part of the literature treats the school (or school district) as the unit of analysis in analyzing outcomes. Available background variables often represent either community characteristics such as measures of community income, racial composition, and so on, or averages of student-body characteristics -- average occupational score and educational score of parents, percent of students who are Black, etc. Outcomes are average performance scores for the school (or for each grade or age group in the school).

Such analyses with average outcome measures at the school level may be useful for performance measurement and goal-setting by schools and school districts, but they have only limited relevance to the objectives of NAEP. The possible impact of some student characteristics on individual student performance tends to get obscured almost completely by this process. For example, Comber and Keeves (7) found a significant sex difference in science scores of high school seniors when the unit of analysis was the individual student, but, at least in public high schools in the United States, the sex distribution tends to be so even that the sex difference would not appear in the differences among schools. Also, racial balancing of schools will tend to obscure racial differences even though those differences may be ascribed to SES or factors other than race itself.

A final problem with using the school as the unit of analysis was pointed out in Section 1.3. When the school is the unit of analysis, the total variance to be explained is the total variance *among* *schools*, which is only one component of total variance among students. Explanation of a certain percentage of between-school variance is not readily interpretable in terms of percent of between-student variance.

However, using school characteristics as the *unit of observation* for defining student independent or background variables, and the student as the *unit of analysis* for outcome variables can be useful both to NAEP and to the more general educational research community. Such uses have been common. The procedure creates an analytical record for each student which contains his individual scores, his individual characteristics and the school (or community) background codes -- percent Black, average income, etc. All students in a given school would have the same coding of school and community background. This mode of analysis was used by Mayeske et al. (31). However, Mayeske's analysis frequently included both the individual student characteristics and the average of them for the student body. As a result, many of the student-body characteristics showed little association beyond that shown by individual characteristics.

A reanalysis of the data could be done to determine percentage of total individual-student variance that could be accounted for by student-body characteristics. Such analyses could be important to NAEP because it is less costly (and sometimes more feasible politically) to obtain some characteristics on a school-wide basis than individually from students. There might be a significant loss in the proportion of the total variance that could be accounted for by such variables unless they are variables that can only be measured at the school level such as school expenditures per student, etc.

A further comment on unit of observation is that none of the large-scale studies we reviewed identified the individual characteristics of the teacher (or teachers) who taught the student whose score was the subject of analysis.[1] Teacher scores (age, race, sex, degrees, scores on word-ability tests, etc.)

---

[1] This may be evidence of the "large-study bias" of the literature reviewed in this study.

were always aggregates (or averages) for the school. Other school experiences of the student might also be related to the individual student. The opinion was expressed in our review panel that school effects might appear more important if the individual teacher, adviser, curriculum, and other specific school inputs for the particular student could have been associated with the individual student scores. It was recognized, however, that the teacher characteristics having the greatest impact on student scores (interest, motivation, "warmth") might be hard to measure, and the student is influenced by teachers he has had in previous years. Also, while it is relatively easy to associate students uniquely with teachers in the lower grades, it becomes increasingly difficult in intermediate and high schools, where there is more assignment of teachers by subject matter. Still, the characteristics of the subject-matter teacher could be associated with the student's scores in that subject matter.

In view of the amount of literature recently that has tended to discount the differential impact of schools and teachers, one can hope that eventually a research project can be mounted to investigate, on a reasonably large scale, the differential impact of teachers and other school inputs on student performance.

5.6 Large-Study Bias

In this study we have intentionally selected the large-scale studies. These have tended to be national or international studies and as a result:

- Certain population subgroups have tended not to receive special emphasis or identification (American Indian, Mexican-American, Chinese, etc.).
- For the most part background characteristics have been selected that are presumed to have an impact for all population subgroups.
- Intensive examination of student-parent-school relationships has not been an important element of most of these studies, although it has received some attention.

These consequences may not be important in meeting the objectives of NAEP, but researchers interested in particular subgroups or in relationships that cannot be obtained easily by standardized tests or questionnaires will not find this study very satisfying.

5.7 Background Adjustment of Longitudinal Data

One of the motivations for adjustment of outcome data for background variables is that the composition of the sample will change over time, reflecting changes in the economic, social, and ethnic characteristics of the population. Without such adjustment an apparent improvement might occur, for example, simply because proportionately fewer educationally disadvantaged persons were included in the population (and in the sample) in the current period than in the base period.

Unfortunately, such adjustment may contain a trap that is not easily recognizable. Suppose, for example, that SES is measured by the number of appliances and educational items in the home (from a standard list of such items). An increase in average score from (say) seven items in the base period to (say) ten items in the current period does not necessarily mean that the SES of students is higher in the current period than in the base period. In part, at least, SES is a relative concept, and as the general level of living increases the number of items in the home tends to increase. Possession of a television set in 1973 hardly has the same meaning as possession of a television set in 1955. The problem is that the base used for adjustment may shift. Even characteristics such as rural, urban, and metropolitan, or geographic division, may lack stability, although perhaps to a considerably lesser extent.

There is no very good solution to this problem. One can, of course, attempt to normalize the scale of the background variables with respect to each year's assessment so that a student's SES score is always relative to the mean (let's say) of that year's assessment group. It seems to us that this may be desirable for some analyses of relative performance rather than absolute performance.

It may be, however, that the variable itself has a different meaning (or impact) over time. For example, the Coleman study was done prior to the extensive efforts to balance school enrollments by race. It seems unlikely that the school characteristic "70 percent Black" has the same meaning now as it had in the middle 1950's. Normalizing the school score relative to the distribution of percent Black is only a partial solution because narrowing the distribution of percent Black by busing and realignment of attendance areas will almost certainly have an impact on the association between outcomes and percent Black in the school.

5.8 Some Additional Comments and Recommendations Concerning Choices of Background Variables for NAEP

a. NAEP has approached the definition of variables used for analysis and adjustment as though the only comparisons to be made are internal comparisons within NAEP, both at one point in time and over time. Thus, NAEP's definition of size and type of community is unique to NAEP and not comparable with classifications used by the Census Bureau or any other agency so far as we are aware. However, NAEP is not action-oriented and one of its purposes is to provide information for general use and evaluation. Consequently, it seems likely that the utility of NAEP results would be enhanced if more attention were given to the choice of classifications so that NAEP results might be compared with information from other sources.

Also, NAEP now collects background information for use as classification or adjustment variables in presenting its test results. We suggest that consideration be given to separate summarization and publication of such background information. This would have the advantage not only of providing socioeconomic measures of direct usefulness in describing the particular populations covered by NAEP, and for measuring changes in these populations, but also for the evaluation of the NAEP sample.

b. We suggest that NAEP should consider introducing an SES measure based on occupation and education of parents for longitudinal comparisons. Items in the home have advantages from the point of view of data collection and should be collected for within-administration analyses. However, as pointed out earlier, items in the home may be less adequate for measuring changes over time, a goal in which NAEP has a primary interest.

Redundancy of information for an SES index appears desirable because of problems of unreliability of response on an individual measure and because of partial nonresponse on tail-sheet (i.e., background) information.

We have given some attention to whether occupational questions should be precoded into a few alternate categories, or whether free answers should be obtained, to be coded subsequently when the returns are received. Contradictory opinions were expressed on this matter in the review of this report by the panel of experts, and we have not seen adequate evidence to make a specific recommendation, especially for the NAEP age groups. We tentatively recommend the check-box approach but urge some additional research directed at this question.

Another approach for an SES index is to obtain address information from the student, code the addresses to Census small areas, and obtain Census community measures of SES that might be assigned on a small-area basis. This approach deserves further exploration. It may be more useful in metropolitan areas than elsewhere.

c. School variables of two types could be obtained by NAEP at modest cost, without substantially increasing the reporting load. School variables have not proved to account for important parts of total variance of individual student achievement-test scores in the studies we have examined, but their contribution has not been inconsequential. Also, there is reason to believe that such variables might be more effective in accounting for change in achievement over time than in accounting for variations in level of achievement at a point in time. We believe this subject deserves attention, and that such variables might be effective for NAEP use. Also, within-school variables, especially characteristics of the teachers of individual students, particular courses taken by the student, grades, and even earlier achievement-test scores might feasibly be collected through the schools, and explored for analytical use. Such information could be obtained without adding to the burden on the student, but the schools might have to be reimbursed for the cost of providing such information.

d. We urge that NAEP consider exploring the utility of a few additional questions on the tail sheet -- possibly questions such as occupational aspirations and future educational intentions. Such questions seem useful within the NAEP environment, and we believe this area deserves more attention and possibly some experimental studies. Another background variable that may be particularly worthwhile exploring is mobility of the student -- possibly whether he lived in the same house, and also in the same county, a year ago, and five years ago.

e. We believe that NAEP can usefully do experimental work of two types: (1) special studies outside of the NAEP operation, and (2) some types of experimental studies incorporated within the ongoing NAEP operation. The latter should be done only after at least limited experimental exploration.

f. We would like to see more attention to evaluation studies than we have seen in NAEP. This is a recommendation we made in an earlier study,[2] and we shall not repeat the recommendations here, but we believe that additional attention in this area would be highly desirable.

g. A literature review such as we have done is rather unrewarding unless it is used to avoid relearning things that have already been learned, or is used to guide future work. We urge, in this connection, that while the literature review can serve these purposes effectively NAEP should not limit consideration of background information to what has been encountered in the literature review. We believe that common sense, judgment, and experimental studies should have a strong additional role in guiding NAEP.

---

[2] "National Assessment Design Implications" by Morris H. Hansen and Edward C. Bryant, presented at the December 1972 meetings of the American Association for the Advancement of Science, Washington, D.C.

5.9 Some Additional Suggestions for Research

This literature review was intended to be one upon which some action can be taken, that is, the choice of a preliminary set of background factors (although feasibility of data collection and advisability of doing so were specifically not a part of the project). We believe that the basic objective of the study has been met, but the study may raise more issues than it resolves. We have made some recommendations here that seem important to us if the association between background factors and educational outcomes is to be better understood and measured.

We suggest two models that may be useful for analytical purposes. Such models would partition the total sum of squares among student scores into the following components separately for White, Black, and other races:

Model I

Among school means		____
Due to school variables	____	
Due to school and community variables	____	
School residual		____
Among students within schools		____
Due to family background	____	
Due to family background and student-parent relationships	____	
Due to family background, student-parent relationships and student affective states	____	
Student residual	____	
Total		____

Model II

Due to family background ______
Due to family background and school ______
Due to family background, school and individual teacher ______
Due to family background, school, teacher and student-parent relationships ______
Due to family background, school, teacher, student-parent relationships and student affective states ______
Residual ______
Total ______

The student is the unit of analysis in both models. The purpose of Model I is analytical, that is, to determine how much of the among-school variance can be accounted for by school and community variables and how much of the within-school variance can be accounted for by student-related characteristics. From this analysis conclusions could be drawn concerning the amount of adjustment one could accomplish by using only school and community characteristics which are inexpensive to gather compared to characteristics of individual students and their families.

Model II orders the analysis in a way that might be useful to NAEP. "Family background" includes SES, family size, education of parents and similar variables. It is that set of variables that is not influenced to any major extent by school policy, funding, or teaching methods. The next variable added is "school," including resources, teaching staff, and similar items. A large increment in explained variance might provide a positive answer to the question: "Do schools make a difference?" Introduction of the individual teachers associated with the student, both in current and prior years, might provide an answer to the question: "Do teachers make a difference?"

Adding the last two variables, student-parent relationships and student affective states have less relevance to near-term educational policy, but the amount of variance explained by them might have relevance to broader-range social policy.

We know of no past study that would provide the basis for complete analysis under these models although the Coleman data, as refined by Mayeske, would provide all but the individual-teacher data.

Secondly, we would like to see an analysis of the percent of variation in student scores that can be accounted for by: (1) individual-student characteristics and family characteristics readily obtainable from students, (2) school characteristics, (3) teacher characteristics (for the specific classroom teacher or the specific subject-matter teacher), and (4) community variables available from secondary sources. An exploratory study could be carried out with and without the more sensitive questions about parents. Such a study would show how much potential for adjustment, if any, is lost by substituting community SES measures for individual-family SES measures.

CITED REFERENCES

CITED REFERENCES

1. Bachman, Jerald G., Robert L. Kahn, Martha T. Mednick, Terrence N. Davidson, and Lloyd D. Johnston, *Youth in Transition: Blueprint for a Longitudinal Study of Adolescent Boys*, Vol. I. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, Mich., 1967.

   Bachman, Jerald G., *Youth in Transition: The Impact of Family Background and Intelligence on Tenth-Grade Boys*, Vol. II. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, Mich., 1970.

2. Bayer, Alan E., Alexander W. Astin, Robert F. Boruch, and John A. Creager, *Users' Manual: ACE Higher Education Data Bank*. ACE Research Reports, Vol. 4, No. 1. Office of Research, American Council on Education, Washington, D.C., 1969.

3. Bayer, Alan E., David E. Drew, Alexander W. Astin, Robert F. Boruch, and John A. Creager, *The First Year of College: A Follow-up Normative Report*. ACE Research Reports, Vol. 5, No. 1. Office of Research, American Council on Education, Washington, D.C., 1970.

4. California State Department of Education, Office of Program Evaluation and Research, *California State Testing Program: 1970-71 Profiles of School District Performance. Technical Supplement*. Sacramento, Calif., 1973.

5. Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York, *Equality of Educational Opportunity*. National Center for Educational Statistics, 1966.

6. Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York, *Supplemental Appendix to the Survey on Equality of Educational Opportunity, Section 9.10: Correlation Tables*. National Center for Educational Statistics, 1966.

7. Comber, L.C. and John P. Keeves, *Science Education in Nineteen Countries*. International Studies in Evaluation I. Stockholm, Sweden: Almquist and Wiksell, 1973.

8. Creager, John A., Alexander W. Astin, Robert F. Boruch, Alan E. Bayer, and David E. Drew, *National Norms for Entering College Freshmen -- Fall 1969*. ACE Research Reports, Vol. 4, No. 7. Office of Research, American Council on Education, Washington, D.C., 1969.

9. Department of Education and Science (Great Britain), *Children and Their Primary Schools: A Report of the Central Advisory Council for Education. Vol. 2: Research and Surveys*. Her Majesty's Stationery Office, London, England, 1967. (The Plowden Report)

10. Donovan, David, Robert Juyser, Philip Kearney, June Olsen, Daniel E. Schooley, and Arthur Carstens, *Distribution of Educational Performance and Related Factors in Michigan*. Sixth Report of the 1970-71 Michigan Educational Assessment Program. Michigan State Department of Education, Research, Evaluation, and Assessment Services, Lansing, Mich., June 1972.

11. Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan, *Socioeconomic Background and Achievement*. New York and London: Seminar Press, 1972.

12. Firman, William D., *The Quality Measurement Project in New York State*. The University of the State of New York, State Education Department, Albany, N.Y. Presented at the American Association for the Advancement of Science meeting in Berkeley, Calif., December 29, 1965.

13. Flanagan, John C., Frederick B. Davis, John T. Dailey, Marion F. Shaycoft, David B. Orr, Isadore Goldberg, and Clinton A. Neyman, Jr., *The American High School Student*. Project Talent Office, University of Pittsburgh, Pittsburgh, Pa., 1964.

14. Flanagan, John C. and William W. Cooley, *Project Talent One-Year Follow-up Studies*. Cooperative Research Project No. 2333. School of Education, University of Pittsburgh, Pittsburgh, Pa., 1966.

15. Fleischmann, Manly, chairman, *The Fleischmann Report on the Quality, Cost, and Financing of Elementary and Secondary Education in New York State*, Vol. I. New York State Commission on the Quality, Cost, and Financing of Elementary and Secondary Education, 1973, New York: The Viking Press, Inc., 1973.

16. Frankel, Martin R., *Inference from Survey Samples: An Empirical Investigation*. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, Mich., 1971.

17. Furno, Orlando F. and George J. Collins, *Class Size and Pupil Learning*. Baltimore City Public Schools, Baltimore, Md., October 1967.

18. Garms, Walter I., *An Approach to the Measurement of Educational Need: The Relationship Between Socioeconomic Characteristics and Pupil Achievement in Basic Skills in Early Elementary School*. Submitted to the New York State Commission on the Quality, Cost, and Financing of Elementary and Secondary Education, August 1971.

19. Hansen, Morris H., William N. Hurwitz, and William G. Madow, *Sample Survey Methods and Theory*, Vol. 1. New York: John Wiley & Sons, 1953.

20. Hanushek, E., *The Value of Teachers in Teaching*. RM-6362-CC/RC. The Rand Corporation, Santa Monica, Calif., December 1970.

21. Hauser, Robert M., "Schools and the Stratification Process," *American Journal of Sociology*, 74 (May 1969), 587-611.

22. Health Services and Mental Health Administration, Division of Emergency Health Services, Department of Health, Education, and Welfare, *School Achievement of Children by Demographic and Socioeconomic Factors, United States*. Washington, D.C., December 1971.

23. Hilton, Thomas L., Albert E. Beaton, and Cathleen Patrick Bower, *Stability and Instability in Academic Growth -- A Compilation of Longitudinal Data*. Educational Testing Service, Princeton, N.J., August 1971. Conducted for the Bureau of Research, Office of Education, Department of Health, Education, and Welfare.

24. Hilton, Thomas L. and Gosta W. Berglund, *Sex Differences in Mathematics -- A Longitudinal Study*. Educational Testing Service, Princeton, N.J., September 1971.

25. Hilton, Thomas L., Haskell Rhett, Irving L. Broudy, Cathleen Bower, Marilyn M. Carter, F. Reid Creech, Gary J. Echternacht, in collaboration with Hunter M. Breland, Franklin R. Evans, and Barbara M. Pinchak, *The Base-Year Survey of the National Longitudinal Study of the High School Class of 1972*. Final Report OEC-0-72-0903. Educational Testing Service, Princeton, N.J., Spring 1972. Conducted for the Office of Education, Department of Health, Education, and Welfare.

26. Husén, Torsten, ed., *International Study of Achievement in Mathematics: A Comparison of Twelve Countries*, Vols. I and II. Stockholm: Almquist and Wiksell and New York: John Wiley & Sons, 1967.

27. Husén, Torsten, *Social Background and Educational Career Research Perspectives on Equality of Educational Opportunity*. Centre for Educational Research and Innovation, Organisation for Economic Cooperation and Development, Paris, France, 1972.

28. Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson, *Inequality -- A Reassessment of the Effect of Family and Schooling in America*. New York, London: Basic Books, Inc., 1972.

29. Kiesling, H.J., *The Relationship of School Inputs to Public School Performance in New York State*. P-4211. Rand Corporation, Santa Monica, Calif., October 1969.

30. Kohen, Andrew I., with the assistance of Paul Andrisani, *Career Thresholds: A Longitudinal Study of the Educational and Labor Market Experience of Male Youth*, Vol. 4. Center for Human Resource Research, Ohio State University, Columbus, Ohio, January 1973.

31. Mayeske, George W., Carl E. Wisler, Albert E. Beaton, Jr., Frederic D. Weinfeld, Wallace M. Cohen, Tetsuo Okada, John M. Proshek, and Kenneth A. Tabler, *A Study of Our Nation's Schools*. Office of Education, Department of Health, Education, and Welfare, Washington, D.C., 1970.

32. Mayeske, George W., Carl E. Wisler, Albert E. Beaton, Jr., Frederic D. Weinfeld, Wallace M. Cohen, Tetsuo Okada, John M. Proshek, and Kenneth A. Tabler, *A Study of Our Nation's Schools*. Office of Education, Department of Health, Education, and Welfare. (Preliminary copy, not published)

33. Mayeske, George W., Tetsuo Okada, Wallace M. Cohen, Albert E. Beaton, Jr., and Carl E. Wisler, *A Study of the Achievement of Our Nation's Students*. DHEW Publication No. (OE) 72-131. Office of Education, Department of Health, Education, and Welfare, Washington, D.C., 1973.

34. Mayeske, George W., Frederic D. Weinfeld, Albert E. Beaton, Jr., Walter David, William B. Fetters, and Eugene E. Hixson, *Item Response Analyses of the Educational Opportunities Survey Student Questionnaires*. Technical Note No. 64. Division of Elementary and Secondary Programs of the Office of Program Planning and Evaluation, and Division of Data Analysis and Dissemination, National Center for Educational Statistics, April 1968.

35. McDonald, Keith H., "The Relationship of Socio-Economic Status to an Objective Measure of Motivation," *Personnel and Guidance Journal*, 42 (June 1964), 997-1002.

36. McPartland, James, "The Relative Influence of School and of Classroom Desegregation on the Academic Achievement of Ninth-Grade Negro Students," *Journal of Social Issues*, 25 (Summer 1969), 93-102.

37. Mosteller, Frederick and Daniel P. Moynihan, eds., *On Equality of Educational Opportunity*. New York: Vantage Books, 1972.

38. Mushkin, Selma J., assisted by Virginia Herman, *The "SIR" Adjusted Index of Educational Achievement*. Public Services Laboratory, Georgetown University, Washington, D.C. Prepared for the Office of Education.

39. Mushkin, Selma J., John Clausen, Harriet Cobern, Morris Cobern, Helen Curtis, Virginia Herman, Wayne Kimmel, James Langenfeld, Margaret Ullman, and Reba V. Wilson, *Educational Outcomes: An Exploratory Review of Concepts and Their Policy Application*. Public Services Laboratory, Georgetown University, Washington, D.C., April 1972.

40. Parnes, Herbert S., Robert C. Miljus, and Ruth S. Spitz and Associates, *Career Thresholds*, Vol. 1. Center for Human Resource Research, Ohio State University, Columbus, Ohio, February 1969. Prepared for the Manpower Administration, U.S. Department of Labor.

41. Purves, Alan C., *Literature Education in Ten Countries*. International Studies in Evaluation II. Stockholm, Sweden: Almquist and Wiksell, 1973.

42. Reiss, Albert J., Jr., with Otis Dudley Duncan, Paul K. Hatt, and Cecil N. North, "A Socioeconomic Index for All Occupations," Chapter VI in *Occupations and Social Status*. The Free Press of Glencoe, Inc., 1961.

43. Research Triangle Institute, Statistics Research Division, *Correlates of Socio-Educational Status CAPE Pilot Study*. Research Triangle Park, N.C., July 1969. Prepared for the Committee on Assessing the Progress of Education.

44. Roderick, Roger D., with the assistance of Joseph M. Davis, *Year for Decision: A Longitudinal Study of the Educational and Labor Market Experience of Young Women*, Vol. 2. Center for Human Resource Research, Ohio State University, Columbus, Ohio, March 1973.

45. Shea, John R., Ruth S. Spitz, and Frederick A. Zeller and Associates, *Dual Careers: A Longitudinal Study of Labor Market Experience of Women*, Vol. 1. Center for Human Resource Research, Ohio State University, Columbus, Ohio, May 1970.

46. Texas Education Agency, *1971 Texas Achievement Appraisal Study*. Austin, Texas, May 1972.

47. Texas Education Agency, *Sixth Grade Reading: A Needs Assessment Report; Sixth Grade Mathematics: A Needs Assessment Report*. Austin, Texas, 1972.

48. Thorndike, Robert L., *Reading Comprehension in Fifteen Countries*. International Studies in Education. New York: John Wiley & Sons, 1973.

49. Tiller , Dale, *Distribution and Differentiation of Youth: A Study of Transition from School to College*. Cambridge, Mass.: Ballinger Publishing Company, 1973.

50. U.S. Department of Health, Education, and Welfare, Bureau of Educational Personnel Development, *Do Teachers Make a Difference: A Report on Recent Research on Pupil Achievement*. Washington, D.C., 1970.

51. University of the State of New York, *Performance Indicators in Education-PIE*. State Education Department, Albany, N.Y., 1972.

52. Wilson, Alan B., *The Consequences of Segregation: Academic Achievement in a Northern Community*. Survey Research Center, University of California, Berkeley, Calif., March 1969.

53. Zeller, Frederick A., John R. Shea, Andrew I. Kohen, and Jack A. Meyer, *Career Thresholds: A Longitudinal Study of the Educational and Labor Market Experience of Male Youth*, Vol. 2. Center for Human Resource Research, Ohio State University, Columbus, Ohio, September 1970. Prepared for the Manpower Administration, U.S. Department of Labor.

BEST COPY AVAILABLE

A MONOGRAPH FROM THE
NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
A PROJECT OF THE EDUCATION COMMISSION OF THE STATES